GENERAL ORACLE INEQUALITIES FOR A PENALIZED LOG-LIKELIHOOD CRITERION BASED ON NON-STATIONARY DATA

BY JULIEN AUBERT[1,a], LUC LEHÉRICY[1,b] AND PATRICIA REYNAUD-BOURET[1,c]

[1]*Université Côte d'Azur, CNRS, LJAD, France,* [a]*julien.aubert@univ-cotedazur.fr;* [b]*luc.lehericy@univ-cotedazur.fr;* [c]*patricia.reynaud-bouret@univ-cotedazur.fr*

We prove oracle inequalities for a penalized log-likelihood criterion that hold even if the data are not independent and not stationary, based on a martingale approach. The assumptions are checked for various contexts: density estimation with independent and identically distributed (i.i.d) data, hidden Markov models, spiking neural networks, adversarial bandits. In each case, we compare our results to the literature, showing that, although we lose some logarithmic factors in the most classical case (i.i.d.), these results are comparable or more general than the existing results in the most dependent cases.

1. Introduction. Maximum likelihood estimator (MLE) (see [13] and references therein) is often considered *the* default approach to construct estimators, although it has been debated, particularly with regard to robustness [8]. In the i.i.d. setting, it is known to be asymptotically efficient (its variance is asymptotically minimal w.r.t. the Cramer-Rao bound) under mild conditions (typically differentiability of the density) [41]. In 1973, Akaike [1] proposed his famous penalized log-likelihood criterion (AIC) stating that to select a model from a finite set of models, one must penalize the log-likelihood of any model m by D_m, the number of parameters describing m. There exists a large variety of variants of AIC, for which the asymptotic properties are more or less precise (see [19] and references therein). Under additional differentiability assumptions, the Wilks phenomenon [55] more precisely quantifies (asymptotically) how the recentered maximal log-likelihood behaves as a chi-square distribution with D_m degrees of freedom. This phenomenon makes it possible to construct asymptotic likelihood ratio tests and therefore to carry out model selection by multiple testing [57]. This idea has been used in many contexts, most recently in combination with asymptotic model selection ℓ_1 [52, 50, 49]. In short, with AIC type penalties and the Wilks phenomenon, there is a fairly clear understanding of how the maximum likelihood behaves asymptotically in the i.i.d. setting in terms of dimension and how the log-likelihood should be penalized to find the correct model in a finite fixed set of smooth enough models.

If we move on to the non-asymptotic framework, things are more difficult to study. There are concentration inequalities that mimic the Wilks phenomenon [16] and provide a non-asymptotic understanding of what the MLE does in an exponential family [47], or how to penalize it so that it works even in infinite-dimensional settings with finite effective dimension [48]. It is even possible to penalize the log-likelihood in order to perform model selection [20, 42] for a particular family of models in the i.i.d. framework. However, in these works, model selection "à la Birgé-Massart" [15] is usually based on some form of linearity between the contrast and the family of models considered, which leads to a very fine tuning of the penalty constants [2] but, at the same time, prevents the results from being applied to more general contexts.

If we drop the independence hypothesis but keep the stationarity of the data, there is a wide variety of model selection results obtained by minimizing penalized contrasts, with

MSC2020 subject classifications: Primary 62G05; secondary 62M99, 62M05.
Keywords and phrases: log-likelihood, penalization, oracle inequality, dependent data.

ℓ_0 or ℓ_1 penalty, with or without log-likelihood. These include Markov chains [37], hidden Markov models [38, 24, 40], spiking neuronal networks with unobserved components [43], point processes [35]. In each case, the arguments are a combination of a martingale approach and non-asymptotic exponential inequalities, which derive from the ergodicity of the process and mixing properties.

However, the Akaike criterion (minus log-likelihood plus penalization proportional to D_m) is often used even outside the frameworks mentioned above [23, 56]. For instance, in learning experiments where an individual has to learn to perform a task and learn it only once, the data are neither independent nor stationary. Some authors [45] have tried to assume that individuals are identically distributed, but this is a very strong and quite unlikely setting in practice. In a first work [5], we proved meaningful bounds for the maximum likelihood estimator on an individual learning trajectory. But, to our knowledge, there is no theoretical work on model selection in this framework or, more generally, for dependent and non-stationary data.

The aim of this work is to derive a non-asymptotic oracle inequality for AIC-type model selection that is general enough to cover all the configurations listed above: i.i.d. samples, hidden Markov models, partially observed neural networks, learning models and more.

The proof relies on an exponential inequality that holds for a supremum of empirical centered processes that are stochastically normalized under a Hölder condition on the parametrization. It is a variant of the Lipschitz case, that we recently proved [4]. This exponential inequality generalizes the works of [7], [51], and is inspired by the works on renormalized martingales due to [10, 12] and [26]. The Hölder case is especially useful to derive bounds for hidden Markov models.

The rest of the article is organized as follows. In Section 2, we introduce notations and the general framework. In Section 3, we state our assumptions and oracle inequalities in probability and in expectation, in the bounded and unbounded frameworks. We also discuss how the penalty in D_m is obtained with respect to more conventional proof techniques. In Section 4, we examine various cases covered by these general results in the light of existing results in the literature. The appendices are devoted to proofs and to the exponential inequality under Hölder parametrization.

2. Framework and notation.

Given two integers $a \leqslant b$ and a sequence $(x_s)_{s\in\mathbb{Z}}$, write $x_a^b = (x_a, \dots, x_b)$ (with x_a^b being the empty sequence when $a > b$). For any two real numbers x, y, write $x \vee y$ their maximum and $x \wedge y$ their minimum. Let $\mathbb{N}^*$ be the set of positive integers, and for any $n \in \mathbb{N}^*$, write $[n]$ the set of integers $\{1, \dots, n\}$. Finally, $\log$ denotes the natural logarithm.

Let $n \geqslant 3$ be an integer. We observe a process $(X_t)_{1\leqslant t\leqslant n}$ defined on a polish measure space $(\mathcal{X}, \mathcal{F}, \mu)$ and adapted to a filtration $(\mathcal{F}_t)_{1\leqslant t\leqslant n}$. Let us write $\mathbb{P}$ the corresponding probability and $\mathbb{E}$ the corresponding expectation. The abbreviation "a.s." is used for "almost surely under the true distribution $\mathbb{P}$" (unless stated otherwise).

The objects of study are the successive conditional distributions of X_t. If $\mathcal{X}$ is discrete and μ is the counting measure, they are defined by the sequence

$$p_t^\star(.) = \mathbb{P}(X_t = .|\mathcal{F}_{t-1}), \quad \forall t \in [n].$$

More generally, in the sequel, we denote by μ a fixed measure on $\mathcal{X}$. For general measured spaces $(\mathcal{X}, \mathcal{F}, \mu)$, we always assume that the conditional density of X_t given $\mathcal{F}_{t-1}$ with respect to μ exists and we denote it by $p_t^\star(.)$. Therefore, for all $x \in \mathcal{X}$, $(p_t^\star(x))_{1\leqslant t\leqslant n}$ is predictable with respect to the filtration—we say that $p_t^\star$ is predictable for short. Let $p^\star = (p_t^\star)_{t\in[n]}$ be the vector of all the successive conditional densities.

2.1. *Some examples.*

Let us give some examples of the filtration $\mathcal{F}_t$.

The filtration depends only on past observations. The emblematic case is when the filtration $\mathcal{F}_t = \sigma(X_1^t)$ for $t \geqslant 1$ and $\mathcal{F}_0$ is the trivial σ-algebra. Here, for $t > 1$, $p_t^\star$ can be written as the conditional density of X_t given X_1^{t-1} and $p_1^\star$ as the density of X_1 (in this case $p_1^\star$ is deterministic). To emphasize this fact, in this case we note $p_t^\star(.|X_1^{t-1})$ instead of just $p_t^\star$. The density of the vector X_1^n with respect to $\mu^{\otimes n}$ is therefore

$$x_1^n \mapsto \prod_{i=1}^n p_t^\star(x_t|X_1^{t-1} = x_1^{t-1}). \tag{1}$$

In the even simpler case that the coordinates of X_1^n are independent, for all $t \geqslant 1$, $p_t^\star$ is the (deterministic) density of X_t w.r.t. to μ.

The filtration depends on past observations and additional covariates. Another example is an enlargement of the previous filtration. For instance, the conditional distribution of X_t can be a function not only of past realizations of X_t, but also of additional covariates: at each step t, the distribution of X_t depends not only on the past but also on an observed variable C_t, so that conditional densities can be written

$$p_t^\star(x_t) = p_t^\star(x_t|X_1^{t-1}, C_1^t). \tag{2}$$

Any decent model of evolution for X_t depends on X_1^{t-1} but also on C_1^t. The natural filtration in this context is $\mathcal{F}_{t-1} = \sigma(X_1^{t-1}, C_1^t)$.

The filtration is not fully observed. Chains with infinite memory, which can model potentially infinite neural networks [30, 31, 32, 43], are another example in which the filtration may not be completely observed. Even if the chain is not observed over the entire network with infinite past, $p_t^\star$ still exists as a function of the infinite network. However $p_t^\star$ is approximated by functions that involve only a smaller subset of the observed neurons (see Section 4.3 for more details).

2.2. *Models and penalized (partial) log-likelihood.* We model $p^\star$ by models that depend on a finite number of parameters. Consider a sequence of models $(\{p_\theta^m : \theta \in \Theta_m \subset \mathbb{R}^{D_m}\})_{m\in\mathcal{M}}$ for some countable set $\mathcal{M}$. Each p_θ^m is a sequence $p_\theta^m = (p_{\theta,t}^m)_{t\in[n]}$, with $p_{\theta,t}^m$ being a candidate at being $p_t^\star$. In particular, the candidate $p_{\theta,t}^m$ must be predictable.

2.2.1. *Partial log-likelihood.* For any $m \in \mathcal{M}$, define the (partial) log-likelihood of parameter $\theta \in \Theta_m$ given the observations by

$$\ell_n(\theta) = \sum_{t=1}^n \log p_{\theta,t}^m(X_t).$$

In the case mentioned in Section 2.1 where $\mathcal{F}_t = \sigma(X_1^t)$ for $t \geqslant 1$ and $\mathcal{F}_0$ is the trivial σ-algebra, because of (1), $\ell_n(\theta)$ is exactly the log-likelihood $\log p_\theta^m(X_1^n)$. For models as in (2), this partial log-likelihood still has the convenient form of a conditional log-likelihood: $\ell_n(\theta) = \log p_\theta^m(X_1^n|C_1^n)$, and it matches Cox's partial likelihood [22]. This might no longer be the case for larger filtrations.

2.2.2. *Penalized log-likelihood criterion.* For each $m \in \mathcal{M}$, define the maximum likelihood estimator of model m by

$$\hat\theta^m \in \underset{\theta\in\Theta_m}{\arg\max}\ \frac{1}{n}\ell_n(\theta).$$

We assume for simplicity in the sequel that the previous estimator always exists, but the following results hold up to an additional error term in ϵ_0 in the oracle inequalities, if we only have:

$$\sup_{\theta \in \Theta_m} \frac{1}{n} \ell_n(\theta) - \epsilon_0 \leqslant \frac{1}{n} \ell_n(\hat{\theta}^m).$$

Finally, take a penalty $\mathrm{pen} : \mathcal{M} \to \mathbb{R}_+$ and select a model $\hat{m}$ that minimizes the penalized log-likelihood:

$$\hat{m} \in \arg\min_{m \in \mathcal{M}} \left(-\frac{1}{n} \ell_n(\hat{\theta}^m) + \mathrm{pen}(m) \right).$$

The penalized likelihood estimator of $p^\star$ is therefore $\tilde{p} = p^{\hat{m}}_{\hat{\theta}^{\hat{m}}}$. We want to understand how good is the estimator $\tilde{p}$ with respect to $p^\star$.

2.3. *Stochastic risk function.* Classical approaches [42, 47, 48], generally use an expectation of the contrast to define the risk. For instance, in i.i.d. examples, the log-likelihood is naturally related to the Kullback-Leibler divergence between the distributions.

Here, because we want to keep the inherent martingale structure that comes with the filtration and with the object of interest $p^\star$, we use the stochastic risk function $\mathbf{K}_n$ defined as follows. For any sequence of conditional densities $p = (p_t)_{t \in [n]}$, let

$$\mathbf{K}_n(p) = \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_t(X_t)} \Big| \mathcal{F}_{t-1} \right].$$

This can be seen as the mean of the conditional Kullback-Leibler divergence in the sense that

$$\mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_t(X_t)} \Big| \mathcal{F}_{t-1} \right]$$

is a predictable quantity which corresponds to the Kullback-Leibler divergence between the distributions with densities $p_t^\star$ and p_t w.r.t μ, conditionally to $\mathcal{F}_{t-1}$.

In the case where $\mathcal{F}_t = \sigma(X_1^t)$ for $t \geqslant 1$ and $\mathcal{F}_0$ is the trivial σ-algebra, because of (1), $n\mathbb{E}[\mathbf{K}_n(p)]$ is exactly the Kullback-Leibler divergence between the distributions defined respectively by $p^\star$ and p.

The oracle inequalities presented below are bounds on the stochastic risk $\mathbf{K}_n(\tilde{p})$ in probability and in expectation respectively.

3. Main results.

3.1. *Main assumptions.* Let us first discuss the main assumptions. Because we use a Kullback-Leibler-like divergence as a risk function, we need to ensure that it does not diverge. A natural assumption is that $p^\star$ and the candidates p^m_θ stay far from 0 and do not explode. This is done almost surely in Assumption 1 and with high probability only in Assumption 1bis.

ASSUMPTION 1. *There exists $\varepsilon > 0$ such that almost surely, for all $t \in [n]$ and $x \in \mathcal{X}$, $p_t^\star(x) \in (\varepsilon, \varepsilon^{-1})$ and for all $m \in \mathcal{M}$ and all $\theta \in \Theta_m$, $p^m_{\theta,t}(x) \in (\varepsilon, \varepsilon^{-1})$. Without loss of generality, we assume that* $\log \varepsilon < -1$.

A weaker version of this assumption is the following.

ASSUMPTION 1bis. *For all $m \in \mathcal{M}$, there exists a finite constant B_m such that almost surely, for all $y \geqslant 1$ and $t \in [n]$,*

$$\mathbb{P}\left(\exists m \in \mathcal{M}, \sup_{\delta,\theta \in \Theta_m \cup \{\star\}} \left|\log \frac{p^m_{\delta,t}(X_t)}{p^m_{\theta,t}(X_t)}\right| \geqslant B_m y \,\middle|\, \mathcal{F}_{t-1}\right) \leqslant e^{-y},$$

with the convention $p^m_\star = p^\star$. Without loss of generality, we assume that $B_m \geqslant 1$ for all $m \in \mathcal{M}$.

The second category of assumptions replaces the exponential family assumption of the classic asymptotic results on MLE. It states that the parameterization of the models is Hölder w.r.t. some norm that is bounded on Θ_m.

ASSUMPTION 2. *For all $m \in \mathcal{M}$, there exist a norm $\|\cdot\|_m$ on $\mathbb{R}^{D_m}$ and finite, positive constants L_m, M_m and $0 < \beta_m$ such that a.s., for all $t \in [n]$ and $x \in \mathcal{X}$, for all $m \in \mathcal{M}$ and all $\delta, \theta \in \Theta_m$,*

$$\left|\log \frac{p^m_{\delta,t}(x)}{p^m_{\theta,t}(x)}\right| < L_m \|\delta - \theta\|_m^{\beta_m}$$

and

$$\|\delta - \theta\|_m \leqslant M_m.$$

Without loss of generality, we assume that $L_m M_m^{\beta_m} \geqslant 1$ for all $m \in \mathcal{M}$.

The case $\beta_m < 1$ corresponds to the classic Hölder condition and the case $\beta_m = 1$ to the Lipschitz one. The case $\beta_m > 1$ on the other hand does not correspond to the classical Hölder definition. In simple cases such as when Θ_m is an interval of $\mathbb{R}$, taking $\beta_m > 1$ entails that the likelihoods are constant, case which is of little interest, but the situation might be different if Θ_m is discrete for instance.

Likewise, a weaker version of this assumption is

ASSUMPTION 2bis. *There exists $\alpha \geqslant 1$ such that for all $m \in \mathcal{M}$, there exist a norm $\|\cdot\|_m$ on $\mathbb{R}^{D_m}$ and finite, positive constants L_m, M_m and β_m such that almost surely, for all $t \in [n]$,*

$$\mathbb{P}\left(\exists m \in \mathcal{M}, \sup_{\delta \neq \theta \in \Theta_m} \frac{\left|\log \frac{p^m_{\delta,t}(X_t)}{p^m_{\theta,t}(X_t)}\right|}{\|\delta - \theta\|_m^{\beta_m}} \geqslant L_m \log n \,\middle|\, \mathcal{F}_{t-1}\right) \leqslant n^{-(1+\alpha)}$$

and

$$\|\delta - \theta\|_m \leqslant M_m.$$

Without loss of generality, we assume that $L_m M_m^{\beta_m} \geqslant 1$ for all $m \in \mathcal{M}$.

3.2. *Measurability assumptions.* This section is concerned with checking the measurability of the quantities involved in our proofs and the previous assumptions. It can be skipped by readers not concerned with this issue.

First, the events in Assumptions 1, 1bis, 2 and 2bis are not necessarily measurable. A simple way to solve this is the following assumption:

ASSUMPTION 3. *For all $x \in \mathcal{X}$, $t \in [n]$ and $m \in \mathcal{M}$, the functions*

$$(\delta,\theta) \in \Theta_m^2 \text{ with } \delta \neq \theta \mapsto \frac{\left|\log \frac{p^m_{\theta,t}(x)}{p^m_{\delta,t}(x)}\right|}{\|\delta-\theta\|_m^{\beta_m}} \quad \text{and} \quad \theta \in \Theta_m \mapsto \left|\log \frac{p^m_{\theta,t}(x)}{p^\star_t(x)}\right|$$

are continuous.

The next potential issue is as follows. In the proof, we define a random distance R_∞ in (21) (Section C.2) as a supremum over $\mathcal{X}$ of functions involving $x \mapsto p^\star_t(x)$ and $p^m_{\theta,t}(x)$. The following assumption ensures that it is defined in a way that it is measurable w.r.t. the filtration.

First, if Assumptions 1 and 2 hold, the following assumption is enough:

ASSUMPTION 4. *There exists a countable dense subset $\mathcal{Q}$ of $\mathcal{X}$ such that almost surely, for all $t \in [n]$, $m \in \mathcal{M}$ and $\theta, \delta \in \Theta_m \cup \{\star\}$,*

$$\sup_{x\in\mathcal{Q}} \left|\log \frac{p^m_{\theta,t}(x)}{p^m_{\delta,t}(x)}\right| = \sup_{x\in\mathcal{X}} \left|\log \frac{p^m_{\theta,t}(x)}{p^m_{\delta,t}(x)}\right|$$

with the convention $p^m_\star = p^\star$.

For instance, this holds for any dense set $\mathcal{Q}$ as soon as the functions $x \mapsto |\log \frac{p^m_{\theta,t}(x)}{p^m_{\delta,t}(x)}|$ are lower semi-continuous for all $\theta, \delta \in \Theta_m \cup \{\star\}$, and it is possible to take $\mathcal{Q}$ countable since $\mathcal{X}$ is separable.

When Assumptions 1bis and 2bis hold, we need a different assumption, such as the following.

ASSUMPTION 4bis. *For all $x \in \mathcal{X}$, there exists an open set $O_x \subset \mathcal{X}$ such that $x \in \overline{O_x}$, the closure of O_x, and for all $t \in [n]$, the set of functions*

$$\left\{ y \in O_x \cup \{x\} \mapsto \frac{\left|\log \frac{p^m_{\theta,t}(y)}{p^m_{\delta,t}(y)}\right|}{L_m\|\delta-\theta\|_m^{\beta_m}},\ m \in \mathcal{M},\ \delta,\theta \in \Theta_m, \delta \neq \theta \right\}$$
$$\cup \left\{ y \in O_x \cup \{x\} \mapsto \frac{1}{B_m}\left|\log \frac{p^m_{\theta,t}(y)}{p^\star_t(y)}\right|,\ m \in \mathcal{M},\ \theta \in \Theta_m \right\}$$

is equicontinuous at x.

For instance, when $\mathcal{X}$ is discrete, both Assumptions 4 and 4bis hold.

For general $\mathcal{X}$, a convenient situation for Assumptions 4 and 4bis is when the densities are piecewise continuous, that is there exists a family of disjoint open sets $(O_i)_{i\in I}$ for an at most countable set I whose closures cover $\mathcal{X}$ and such that the restriction of the densities on each of these sets can be extended into a continuous function over $\mathcal{X}$. We actually need something a bit stronger than piecewise continuous to deal with the possible discontinuity at the border of each open set: Assumption 4bis holds as soon as there exists a partition $(A_i)_{i\in I}$ of $\mathcal{X}$ with $O_i \subset A_i \subset \overline{O_i}$ (for the open sets O_i discussed above) such that the functions $y \mapsto \left|\log \frac{p^m_{\theta,t}(y)}{p^m_{\delta,t}(y)}\right| / (L_m\|\delta-\theta\|_m^{\beta_m})$ and $y \mapsto \left|\log \frac{p^m_{\theta,t}(y)}{p^\star_t(y)})\right| / B_m$ (for all m and δ, θ) are equicontinuous on each A_i. The partition $(A_i)_{i\in I}$ must be the same for all densities of all models.

EXAMPLE 1. Take $\mathcal{X} = \mathbb{R}$, $n = 1$ and the single model $\Theta = \{p_\lambda : x \mapsto \lambda e^{-\lambda x}\mathbf{1}_{x \geqslant 0}, \lambda \in [a,b]\}$ with $a, b \in \mathbb{R}_+^*$, and assume $y \mapsto p^\star(y)$ is zero on $\mathbb{R}_-^*$ and continuous and positive on $\mathbb{R}_+$. These densities satisfy the condition discussed above with $O_1 = \mathbb{R}_-^*$, $O_2 = \mathbb{R}_+^*$, $A_1 = \mathbb{R}_-^*$ and $A_2 = \mathbb{R}_+$, with Hölder regularity $\beta = 1$.

When $\mathcal{M}$ is finite, the following assumption suffices and is typically easier to check.

ASSUMPTION 4ter. *The set $\mathcal{M}$ is finite and for all $x \in \mathcal{X}$, there exists an open set $O_x \subset \mathcal{X}$ such that $x \in \overline{O_x}$ and for all $t \in [n]$ and $m \in \mathcal{M}$, the functions*

$$(\delta, \theta, y) \in \Theta_m^2 \times (O_x \cup \{x\}) \mapsto \frac{\left|\log \frac{p_{\theta,t}^m(y)}{p_{\delta,t}^m(y)}\right|}{\|\delta - \theta\|_m^{\beta_m}} \quad \textit{and} \quad (\theta, y) \in \Theta_m \times (O_x \cup \{x\}) \mapsto \left|\log \frac{p_{\theta,t}^m(y)}{p_t^\star(y)}\right|$$

are continuous at (δ', θ', x) (with a continuous extension when $\delta' = \theta'$) and (θ', x) respectively, for all $\delta', \theta' \in \Theta_m$.

This assumption implies both Assumptions 3 and 4bis.

3.3. *Oracle inequality, bounded case.*

THEOREM 2. *Assume that $n \geqslant 3$ and that Assumptions 1, 2, 3 and 4 hold. For each $m \in \mathcal{M}$, let $A_m = L_m M_m^{\beta_m} + 2\log(\varepsilon^{-1})$, and assume that*

$$\Sigma = \sum_{m \in \mathcal{M}} \log(A_m) e^{-D_m} < +\infty.$$

Let $\kappa \in (0,1]$. There exist positive numerical constants C and C' such that the following holds. If for all $m \in \mathcal{M}$,

$$\mathrm{pen}(m) \geqslant \frac{C}{\kappa} A_m^2 \log(\varepsilon^{-1}) \left(1 + \frac{1}{\beta_m}\right) \log(n A_m)^2 \frac{D_m}{n},$$

then for all $x \geqslant 0$, with probability at least $1 - 24\log(n)\Sigma e^{-x}$,

$$(1-\kappa)\mathbf{K}_n(\tilde{p}) \leqslant \inf_{m \in \mathcal{M}} \left((1+\kappa) \inf_{\theta \in \Theta^{D_m}} \mathbf{K}_n(p_\theta^m) + 2\mathrm{pen}(m) + \frac{C'}{\kappa}(A_m + A_{\hat{m}}) \log(\varepsilon^{-1}) \frac{x}{n} \right).$$

PROOF. See Section A. □

This result is an oracle inequality in probability. The penalty term is proportional to D_m/n, as in classical oracle inequalities for nested or not too complex families of models (e.g. Gaussian model selection [15]). This is ensured by the summability condition on Σ (see [15, 42] for instance for a discussion about the complexity of a family of models). The risk is—up to a constant factor and a residual term—smaller than the best bias-variance trade-off in the family of models, with a variance of order D_m/n.

Due to the generality of the result, the penalty is a bit larger than in the original AIC criterion, with additional logarithmic factors, and depends on the lower bound ε and the Hölder constants L_m and M_m. Because $\mathcal{M}$ might be infinite, there is a residual term depending on $\hat{m}$. Additional assumptions are required to get rid of it, such as a uniform bound on $(A_m)_{m \in \mathcal{M}}$ introduced in the following corollary which gives a result in expectation and whose proof can be found in Section B.

COROLLARY 3. *Under the same assumptions and with the constants and notations of Theorem 2, if there exists $A(n) > 0$ such that*

$$\sup_{m\in\mathcal{M}} A_m \leqslant A(n),$$

then

$$(1-\kappa)\mathbb{E}\left[\mathbf{K}_n(\tilde{p})\right] \leqslant \mathbb{E}\left[\inf_{m\in\mathcal{M}}\left((1+\kappa)\inf_{\theta\in\Theta^{D_m}}\mathbf{K}_n(p_\theta^m)+2\mathrm{pen}(m)\right)\right] + \frac{36C'}{\kappa}\Sigma A(n)\log(\varepsilon^{-1})\frac{\log n}{n}.$$

3.4. *Oracle inequality, unbounded case.* The bounded case can be restrictive for some applications, so we can relax the assumptions, up to additional logarithmic factors.

THEOREM 4. *Assume that $n \geqslant 3$ and that Assumptions 1bis, 2bis, 3 and 4bis hold. For each $m \in \mathcal{M}$, let $A_m = L_m M_m^{\beta_m} + (1+\alpha)B_m$, and assume that*

$$\Sigma = \sum_{m\in\mathcal{M}} \log(A_m)e^{-D_m} < +\infty$$

Let $\kappa \in (0,1]$. There exist positive constants C and C' such that the following holds. If for all $m \in \mathcal{M}$,

$$\mathrm{pen}(m) \geqslant \frac{(1+\alpha)C}{\kappa}A_m^2 B_m(\log n)^3\left(1+\frac{1}{\beta_m}\right)\log(nA_m)^2\frac{D_m}{n}$$

then for all $x \geqslant 0$, with probability at least $1-2n^{-\alpha}-24\log(n)\Sigma e^{-x}$,

$$(1-\kappa)\mathbf{K}_n(\tilde{p}) \leqslant \inf_{m\in\mathcal{M}}\left((1+\kappa)\inf_{\theta\in\Theta^{D_m}}\mathbf{K}_n(p_\theta^m)+2\mathrm{pen}(m)+\frac{(1+\alpha)C'}{\kappa}(A_mB_m+A_{\hat{m}}B_{\hat{m}})\frac{x(\log n)^2}{n}\right).$$

PROOF. See Section A. □

Except for extra logarithmic factors, Theorem 4 is essentially the same as Theorem 2. An expectation version of this result holds under the same assumptions as for Corollary 3. Its proof can be found in Section B.

COROLLARY 5. *Under the assumptions and with the same constants and notations of Theorem 4, if $\alpha \geqslant 1$ and there exist $A(n) > 0$ and $B(n) > 0$ such that*

$$\sup_{m\in\mathcal{M}} A_m \leqslant A(n) \quad \text{and} \quad \sup_{m\in\mathcal{M}} B_m \leqslant B(n),$$

then

$$(1-\kappa)\mathbb{E}\left[\mathbf{K}_n(\tilde{p})\right] \leqslant \mathbb{E}\left[\inf_{m\in\mathcal{M}}\left((1+\kappa)\inf_{\theta\in\Theta^{D_m}}\mathbf{K}_n(p_\theta^m)+2\mathrm{pen}(m)\right)\right] + \frac{52(1+\alpha)C'}{\kappa}\Sigma A(n)B(n)\frac{(\log n)^3}{n}.$$

3.5. *Origin of the penalty in D_m/n.* The common crucial point to all non-asymptotic controls of log-likelihood estimators [42, 20, 47, 48] lies in the control of the recentered contrast at the estimation point. In our framework, with the above notation, it means controlling

$$\nu(H) = \frac{1}{n}\sum_{t=1}^{n}[H_t(X_t) - \mathbb{E}(H_t(X_t)|\mathcal{F}_{t-1})],$$

where H_t is equal to

$$H_t = H^m_{\theta,t} = -\log\left(\frac{p^m_{\theta,t}(X_t)}{p^\star_t(X_t)}\right).$$

As long as θ is fixed and deterministic, this is a martingale and various exponential tail bounds are applicable. For now, without going into details, let us just say that with high probability,

$$\nu(H) = \mathcal{O}\left(\sqrt{V(H)/n}\right) \tag{3}$$

where typically,

$$V(H) = \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}(H_t(X_t)^2|\mathcal{F}_{t-1}).$$

In the i.i.d. framework, where $V(H)$ is deterministic, of the order of the variance of H_1, we recover what the central limit theorem implies, i.e. that the fluctuations of the empirical process are of order $n^{-1/2}$ multiplied by the standard deviation. In more general settings, $V(H)$ is random and it is usually necessary to restrict oneself to an event where $V(H)$ is bounded to obtain such non-asymptotic control of $\nu(H)$ (see for instance [11]).

However, this is still not sufficient to conclude. To understand what happens for the maximum likelihood estimator, we need to control $\nu(H^m_{\hat\theta_m,t})$ and because $\hat\theta_m$ depends on the whole trajectory, the classical inequalities for centered processes do not apply. In addition to that, the order of magnitude of $V(H)$ in $n^{-1/2}$ is not even the good order of magnitude for the penalty.

3.5.1. *Talagrand inequality and Wilks phenomenon in the i.i.d. setting.* In the classical i.i.d. setting, and as a first approach, Talagrand's inequality (see [42] for various model selection contexts) leads to this kind of control:

$$\sup_{\theta\in\Theta_m}\frac{\nu(H^m_\theta)}{\|\theta\|} = \mathcal{O}\left(\mathbb{E}\left(\sup_{\theta\in\Theta_m,\|\theta\|\leqslant 1}\nu(H^m_\theta)\right)\right) + \mathcal{O}\left(\sqrt{\frac{\sup_{\theta\in\Theta_m,\|\theta\|\leqslant 1}V(H^m_\theta)}{n}}\right).$$

It turns out that in many models used in [42], $\nu(H^m_\theta)$ is linear or close to linear with respect to θ, so that the first term is of order $\sqrt{D_m/n}$, the dimension of the model m. As a result, non-asymptotically, it holds that

$$\nu(H^m_{\hat\theta_m}) = \mathcal{O}\left(\|\hat\theta_m\|\sqrt{D_m/n}\right).$$

Beyond the i.i.d. setting, similar results can be obtained by replacing Talagrand's inequality with Baraud's inequality [7].

In comparison, the Wilks phenomenon predicts that the order of magnitude of $\nu(H^m_{\hat\theta_m})$ is in D_m/n, and not $\sqrt{D_m/n}$ (at least if the model m is well specified, that is there exists $\theta^\star_m$ such that $p^\star = p^m_{\theta^\star_m}$): this non-asymptotic Talagrand-like bound is pessimistic. But the trick above with linearity with respect to θ shows that this pessimistic bound can be multiplied by the norm of $\hat\theta_m$. In itself, this is not much, but by carefully choosing the set over which the supremum is taken, $\|\hat{\theta}_m\|$ can be replaced by the distance to a pivot $\theta^\star_m$ such that, approximately,

$$\nu(H^m_{\hat{\theta}_m}) = \mathcal{O}\left(\|\hat{\theta}_m - \theta^\star_m\|\sqrt{D_m/n}\right) \leqslant \mathcal{O}\left(\|\hat{\theta}_m - \theta^\star_m\|^2\right) + \mathcal{O}\left(\frac{D_m}{n}\right). \tag{4}$$

It remains to be in a sufficiently convenient setting (such as the Hölder parameterization of $\theta \mapsto p^m_\theta$) for $\mathcal{O}\left(\|\hat{\theta}_m - \theta^\star_m\|^2\right)$ to be a small fraction of the risk we are using (here $\mathbf{K}_n(p^m_{\hat{\theta}_m})$). Overall, up to a constant, the remainder that the penalty is supposed to take into account is indeed in D_m/n up to multiplicative constants.

Before going further, let us make a remark about the pivot $\theta^\star_m$. If the model m is well specified, it is easy to use the $\theta^\star_m$ such that $p^\star = p^m_{\theta^\star_m}$. If this is not the case, people generally use the best approximation of $p^\star$ by a p^m_θ for a certain risk function (say the Kullback-Leibler divergence). So, even in non-i.i.d. cases, the classical method of choosing a pivot is deterministic, because the risk is too.

3.5.2. *Non-asymptotic Wilks-like results.* In more complex models, the trick given above (4) cannot work as well because it relies mainly on linearity. The brute force concentration inequality "à la Talagrand" becomes too pessimistic. To obtain similar results, we would like, in an ideal, over-simplified world, for (3) to hold with a variance term that would be directly $V(H^m_{\hat{\theta}_m})$, without having to take (and pay for) the supremum over all H^m_θ.

In [16], authors restrict the previous supremum to nice small balls, so that one gets the correct behavior in D_m/n for $\nu(H^m_{\hat{\theta}_m})$ directly. This can be useful when one wants to produce sharp constants in the penalty [3].

Another way to think about this, is to provide an upper function of the process, that is to prove that

$$\sup_{\theta\in\Theta_m}\left[\nu(H^m_\theta) - \mathcal{O}\left(\sqrt{V(H^m_\theta)/n}\right)\right]$$

is negative with high probability. This is Spokoiny's approach [47, 48], whose statistical results are closest to ours for their generality, even though the author does not perform model selection per se.

3.5.3. *Self-normalized martingales.* It is far from obvious to obtain an inequality such as (3) when $V(H)$ is random. It is possible for martingales, in which case $V(H)$ is usually a random quantity called the bracket of the martingale. Exponential inequalities for martingales have been developed. For instance, for point processes, one can move from a control of the martingale with a deterministic upper bound on the bracket, which already tells a lot on the properties of the maximum likelihood estimator [53], to a control where the bracket of the martingale can replace the deterministic upper bound in the deviation [35] (up to some corrections).

In the same line, many works on self-normalization of martingales try to directly control the ratio of the martingale by its bracket [11, 27]. To our knowledge, nothing exists in this direction for a supremum of the ratios (with random renormalization) except the recent result [4] and its extension to the Hölder parametrization that we use here.

3.5.4. *Deterministically renormalized empirical process.* From a more deterministic point of view, several works aim to choose the correct deterministic renormalization of the empirical process in an i.i.d. setting. There are two main ways in which this can be used for model selection.

On the one hand, if the empirical process itself, once renormalized by a deterministic quantity of the form $d^2(\theta, \theta^\star)$ [1], satisfies a convenient exponential inequality, then this can be chained to either directly obtain a Talagrand-like concentration on the supremum [7], or used to get nice upper-functions [47].

On the other hand, one can refine the renormalization inside the supremum and replace

$$\sup_{\theta\in\Theta_m} \frac{\nu(H_\theta^m)}{\|\theta\|} \quad \text{by} \quad \sup_{\theta\in\Theta_m} \frac{\nu(H_\theta^m)}{d^2(\theta,\theta^\star)+x^2}, \tag{5}$$

for a positive real number x to be chosen later. With the latter, Massart [42] proposed a fairly general approach to model selection with penalty proportional to D_m even in non linear settings. Here, the order of magnitude of the Wilks phenomenon cannot be obtained directly, but since, as before, $d^2(\theta, \theta^\star)$ is close to the risk, it is still possible to get a penalty in D_m/n. The main advantage is the possibility to move away from the linearity assumption, adding more flexibility to the family of models by using $d^2(.,.)$ instead of the Euclidean norm.

3.5.5. *Our method and the difference with existing works.* In our case, even the risk is stochastic in the most general case. We do not have access to a deterministic $d^2(.,.)$ and we cannot properly define a deterministic pivot $\theta_m^\star$ if the model is misspecified. Only the martingale structure and the bracket of the martingale $V(H_\theta)$ remain. In fact, $V(H_{\hat\theta_m}^m)$ is conveniently comparable to $\mathbf{K}_n(p_{\hat\theta_m}^m)$. The recent concentration inequality for the supremum of stochastically normalized processes [4] and the extension we derive here in the appendix makes it possible to use an argument similar to [42] after replacing $d^2(.,.)$ with the bracket of the martingales in (5). The problem of the pivot is solved by working directly in the probability space and using $p^\star$ as pivot, thus bypassing the misspecification issue.

Let us now compare our set of assumptions to Spokoiny's work [47, 48] which constitute, to our knowledge, the most general non-asymptotic results on maximum likelihood estimation. Spokoiny works with the Kullback-Leibler divergence as a reference. He uses a deterministic pivot defined as the closest point to the truth inside the model for the deterministic distance [47] or with additional quadratic corrections [48]. His main assumptions rely on renormalized exponential inequalities on the gradient of the likelihood in a neighborhood of the pivot, where the normalization is quadratic. He proves a quadratic-like behavior for $\nu(H_{\hat\theta_m}^m)$, demonstrating the non-asymptotic Wilks phenomenon rather sharply in this sense, which we cannot do with our method.

On the other hand, our method is applicable to more general settings: we do not need to use a deterministic risk or deterministic parametric pivots, nor do we need to assume that the log-likelihood is differentiable or that its gradient satisfies exponential inequalities. We only assume that the parameterization is Hölder. Thanks to the properties of martingales, the log-likelihood (and not its gradient) automatically satisfies (3), which is the key to proving the results on the supremum. In the end, we obtain a generalized AIC criterion (that is, a penalty proportional to D_m/n) and prove non-asymptotic oracle inequalities in settings where none of the existing work applies.

4. Applications. The preceding oracle inequalities are very general. The aim of this section is to explain how they compare with existing results. Let us look at their applications in various contexts.

[1]for a nice deterministic d distance on Θ_m, that can be linked to $V(H_\theta^m)$

4.1. *The i.i.d. sample case.* Let us begin with the original AIC setting in its i.i.d. format. Let $X_1, \dots, X_n$ be i.i.d. real valued random variables with density $p_1^\star$. Consider the simple case where the filtration is generated by the observations, so that all the $p_t^\star$ are deterministic and equal to $p_1^\star$.

Assume that under each model, the variables $X_1, \dots, X_n$ are also modeled as i.i.d., so that $p_{\theta,t}^m = p_{\theta,1}^m$ for all t and these functions are deterministic.

In this case, $\mathbf{K}_n(p_\theta^m)$ is directly $\mathrm{KL}(p_1^\star d\mu, p_{\theta,1}^m d\mu)$, the Kullback-Leibler divergence between the distribution of X_1 and the distribution with density $p_{\theta,1}^m$.

4.1.1. *Validation of the assumptions.* Assumption 1 is classic in this setting, see e.g. [42], at least for the lower bound. In [42] or [20], there is no upper bound assumption but there is a difference: the Hellinger distance is controlled instead of the Kullback-Leibler divergence. The lower bound assumption can also be relaxed by using Assumption 1bis.

Assumption 2—the L_m part—can be a consequence of the fact that $p_{\theta,1}^m$ is Hölder with constant $L_m\varepsilon^{-1}$ and exponent β_m, or of directly assuming that $\log(p_{\theta,1}^m)$ is Hölder with constant L_m and exponent β_m. It can be relaxed with Assumption 2bis. Classical asymptotic results (Wilks phenomenon or even just the consistency of the MLE) require very strong differentiability assumptions, that are not needed here.

The bound M_m in Assumptions 2 and 2bis entails that the models Θ_m are compact, which is assumed in general to obtain consistency in M-estimation, whether explicitely or implicitely by assuming that the estimator converges to some limit, see e.g. [54].

The classical model selection "à la Birgé-Massart" for densities assumes that the models are close to linear, which might seem at first glance incompatible with our boundedness assumption [42]. However this is not the case because the parameterization is forced to be a density that satisfies Assumption 1. To illustrate this fact and compare our assumptions with theirs, let us restrict ourselves to a well known case: the histogram selection [42, 20]. In this setup, $X_1, \dots, X_n$ are i.i.d. with density $p_1^\star$ with respect to the Lebesgue measure on $[0,1]$ and the model m is based on a partition with D_m intervals of $[0,1]$ of equal length. Then, for $\theta \in \Theta_m \subset \mathbb{R}^{D_m}$,

$$p_{\theta,1}^m = \sum_{I \in m} \theta_I \mathbf{1}_I.$$

Since this must be a density, the model is in fact

$$\Theta_m = \left\{ \theta = (\theta_I)_{I\in m} \text{ such that for all } I, \varepsilon \leqslant \theta_I \leqslant \varepsilon^{-1} \text{ and } \sum_{I\in m} \frac{\theta_I}{D_m} = 1 \right\}.$$

In this particular case, $p_{\theta,1}^m$ is Lipschitz with constant $L_m = 1$ w.r.t. $\|\theta\|_\infty$ and $M_m = \varepsilon^{-1} - \varepsilon$ is an upper bound of the diameter of Θ_m for this norm.

Assumption 3 is straightforward. Assumption 4 is automatically fulfilled with $\mathcal{Q} = \mathbb{Q} \cap [0,1]$, the set of rational numbers, or any other dense countable subset of $[0,1]$.

4.1.2. *Result.* Following Theorem 2 and its Corollary, A_m does not depend on m anymore. Therefore, our oracle inequality holds as soon as

$$pen(m) = \mathcal{O}\left((\log n)^2 \frac{D_m}{n} \right).$$

Our penalty is larger than the one in [42] due to extra logarithmic factors and looks more like a BIC criterion. At this price, we prove an oracle inequality directly on the Kullback-Leibler divergence, instead of a mixed oracle inequality involving both the Kullback-Leibler divergence and the Hellinger distance.

4.2. *Hidden Markov models.* A hidden Markov model is a stochastic process $(S_t, X_t)_t$ where only the observations $(X_t)_t$ are observed, such that the hidden process $(S_t)_t$ is a Markov chain and such that all the X_s's are independent conditionally to $(S_t)_t$ with a distribution depending only on the corresponding S_s. These models have been widely used since their formalization by [9], as they are able to account for complex dependencies in time processes while keeping a very simple and easily interpretable structure.

In this section, we observe the process $(X_t)_t$ and want to approximate $p_t^\star(X_t)$ as before, where the filtration $\mathcal{F}_t$ is the one generated by the X_t's. The process X_t takes values in a measured Polish space $(\mathcal{X}, \mu)$. The hidden Markov models are a family of models used to approximate $p^\star$, we do not assume that $p^\star$ itself is a hidden Markov model: this is what is called a misspecified HMM [40]. We only consider finite state space hidden Markov models, that is $S_t \in [h]$ for some $h \in \mathbb{N}^*$. In particular, the number of hidden states h is a parameter of the models.

The parameters of a hidden Markov model of index m are the number of hidden states h^m, the initial distribution π^m, the transition kernel Q^m of the hidden process $(S_t)_t$ (on $[h^m]$) and the emission densities, that is the family $\nu^m = (\nu_s^m)_{s\in[h_m]}$, where ν_s^m is the density of the distribution of X_1 conditionally to $S_1 = s$ w.r.t. the dominating measure μ.

4.2.1. *Validation of the assumptions.* In what follows, we only assume that the models are HMM, while the true distribution of $(X_t)_t$ may not be one. As such, we treat the true distribution separately, before introducing the models.

First, assume that $p_t^\star$ is continuous and positive; this will be used to check Assumption 4bis.

Concerning Assumption 1bis for the true distribution, the lower tails of $\log p_t^\star(X_t)$ are always automatically sub-exponential by direct application of Markov's inequality. The control of the upper tails follows from the assumption that the conditional densities of $(X_t)_t$ admit a finite moment, that is, there exist constants $\delta > 0$ and $M_\delta > 0$ such that a.s.,

$$\sup_{t\in[n]} \mathbb{E}[p_t^\star(X_t)^\delta \,|\, \mathcal{F}_{t-1}] \leqslant M_\delta < +\infty. \tag{6}$$

Under this assumption, there exists $B^\star > 0$ depending on δ and M_δ such that for all $y \geqslant 1$ and $t \in [n]$,

$$\mathbb{P}(|\log p_t^\star(X_t)| \geqslant B^\star y \,|\, \mathcal{F}_{t-1}) \leqslant e^{-y}.$$

Let us now introduce the models. Let $C_Q > 0$ and $\alpha \geqslant 1$. For all m, let $G_m = \{g_\eta, \eta \in E_m \subset \mathbb{R}^{e_m}\}$ be a parametric set of probability densities on $\mathcal{X}$ satisfying the following assumption.

ASSUMPTION 4-HMM. *For all $m \in \mathcal{M}$, writing $\|\cdot\|_\infty$ the supremum norm on $\mathbb{R}^{e_m}$, there exists $L_m^g, C_m^g > 0$ such that for all $m \in \mathcal{M}$, all $\gamma, \eta \in E_m$ and all $x \in \mathcal{X}$,*

$$|\log g_\gamma(x) - \log g_\eta(x)| \leqslant L_m^g \|\gamma - \eta\|_\infty,$$

$$\|\gamma - \eta\|_\infty \leqslant 1$$

and

$$(C_m^g)^{-1} < g_\eta(x) < C_m^g.$$

Moreover, the map

$$(\gamma, \eta, x) \in E_m^2 \times \mathcal{X} \text{ with } \gamma \neq \eta \mapsto \frac{g_\gamma(x) - g_\eta(x)}{\gamma - \eta}$$

is continuous and can be extended into a continuous function over $E_m^2 \times \mathcal{X}$.

Let $C_Q \geqslant 1$. The model Θ_m is defined as the set

$$\Theta_m = \Big\{ (\pi_\theta^m, Q_\theta^m, (\eta_i^\theta)_{i\in[h_m]}) \in \mathbb{R}^{h_m} \times \mathbb{R}^{h_m \times h_m} \times (E_m)^{h_m}$$
$$\text{s.t. } \pi_\theta^m \text{ probability vector, } Q_\theta^m \text{ transition matrix and}$$
$$(C_Q)^{-1} \leqslant h_m \pi_\theta^m(i) \leqslant C_Q \quad \text{for all } i \in [h_m],$$
$$(C_Q)^{-1} \leqslant h_m Q_\theta^m(i,j) \leqslant C_Q \quad \text{for all } i,j \in [h_m],$$
$$\text{the emission densities are } \nu_{\theta,i}^m = g_{\eta_i^\theta} \in G_m \text{ for } i \in [h_m] \Big\}.$$

This model is of dimension $D_m = h_m e_m + h_m^2 - 1$. By abuse of notations, and only for the HMM example, we use the notation p_θ^m for not only the conditional densities of X_t given X_1^{t-1} as before, but also all other possible densities, even those involving the hidden variables; which (conditional) density is used will be clear from its arguments. For instance, the likelihood of the observations X_1^n under the parameter $\theta \in \Theta_m$ is

$$p_\theta^m(X_1^n) = \sum_{(i_1,\dots,i_n)\in[h_m]^n} \pi_\theta^m(i_1)\nu_{\theta,i_1}^m(X_1) \prod_{t=2}^n Q_\theta^m(i_{t-1},i_t)\nu_{\theta,i_t}^m(X_t).$$

We endow Θ_m with the norm

$$\|\theta - \delta\|_m = \max(h_m\|Q_\theta^m - Q_\delta^m\|_\infty, h_m\|\pi_\theta^m - \pi_\delta^m\|_\infty, L_m^g \max_i \|\eta_i^\theta - \eta_i^\delta\|_\infty),$$

where L_m^g is defined in Assumption 4-HMM. Note that these norms of differences are always smaller than $M_m = 2\max(L_m^g, C_Q)$.

We assume that $\mathcal{M}$ is finite. In practice, $\mathcal{M}$ is allowed to grow with n, so this is not much of a restriction. Recall that $p_t^\star$ is assumed continuous and positive. Given that for all t, $p_{\theta,t}^m(x)$ is a $\mathcal{C}^\infty$ function of π_θ^m, Q_θ^m, $\nu_{\theta,i}^m(x)$ and $\nu_{\theta,i}^m(X_s)$ for all i and $s < t$, the last part of Assumption 4-HMM implies that Assumption 4ter holds with $\beta_m = 1$ and thus for all $\beta_m \leqslant 1$, thus ensuring Assumptions 3 and 4bis.

Given the upper and lower bounds on the initial distribution and transition matrix, Assumptions 1 and 1bis can be replaced by equivalent assumptions on the average $\bar\nu_\theta^m = \frac{1}{h_m}\sum_{i\in[h_m]} \nu_{\theta,i}^m$ of the emission densities $(\nu_{\theta,i}^m)_{i\in[h_m]}$, since

$$\begin{aligned} p_{\theta,t}^m(X_t) &= \sum_{i,i'\in[h_m]} p_\theta^m(S_{t-1} = i | X_1^{t-1}) Q_\theta^m(i,i')\nu_{\theta,i'}^m(X_t) \\ &\in \left[(C_Q)^{-1}\bar\nu_\theta^m(X_t), C_Q\bar\nu_\theta^m(X_t)\right], \end{aligned}$$

and under Assumption 4-HMM, this show that $p_{\theta,t}^m(X_t) \in (\varepsilon, \varepsilon^{-1})$ a.s., with $\varepsilon = (C_Q C_m^g)^{-1}$. Assumption 1bis follows with $B_m = B^\star \vee (2\log(C_Q C_m^g))$. Note that the filtering distribution $p_\theta^m(S_{t-1} = i|X_1^{t-1})$ can be computed using the Forward algorithm [44], although it is not required to know it to reduce checking Assumption 1bis to the control of $\bar\nu_\theta^m(X_t)$.

While the log-densities at time t are Lipschitz in the parameter, their Lipschitz constant grows exponentially with t, which makes the terms in the oracle inequality explode as n grows. The following Lemma, proved in Appendix D, establishes a Hölder bound that holds uniformly for all t, with a Hölder constant depending on the bound on the transition matrix (and initial distribution) of the Markov chain.

LEMMA 6. *Assume Assumption 4-HMM holds, then Assumption 2 holds with* $(\beta_m)^{-1} = 1 + \frac{\log(2C_Q^2)}{-\log(1-(C_Q)^{-2})} \in (0,1]$, $M_m = 2(L_m^g \vee C_Q)$, *and* $L_m \log n = (112 C_Q^8) \vee (3\log C_m^g)$.

To prove Lemma 6, we approximate p_θ^m by its Markov version of order k, which is Lipschitz with a constant growing exponentially with k, and the Hölder regularity follows from a trade-off on k. This is the main reason for introducing Hölder conditions in the main assumptions instead of only Lipschitz ones.

4.2.2. *Result.* The closest result to ours in this setting is the one from [40], which proves an oracle inequality for a penalized maximum likelihood estimator on misspecified hidden Markov models. Assumptions and proofs are similar to ours, although with a slightly different risk, and the author needs additional assumptions on the true distribution and relies on tools specific to hidden Markov models to obtain the oracle inequality.

To get an order of magnitude of the terms in the oracle inequality, let us take comparable quantities to [40] in the assumptions: up to multiplicative constants, $C_Q = \log n$, $\log C_m^g = \log n$, $L_m^g = \log n$, and thus $B_m = \log n$, $L_m = (\log n)^7$, $M_m = \log n$, $\beta_m = ((\log n)^2 \log\log n)^{-1}$, and thus $A_m \leqslant (\log n)^7$ and $\Sigma = \log\log n$, still up to multiplicative constants.

With these quantities, Corollary 5 results in a penalty of order $(\log n)^{22} \log\log(n) D_m/n$ and an error term of order $(\log n)^{11} \log\log(n)/n$, compared to a penalty in $(\log n)^{17} \log\log(n) D_m/n$ and an error term in $(\log n)^{10}/n$ in [40]. Up to polylog factors, this is the expected order of magnitude for the penalty and residual.

Let us go into more details about the assumptions. First, [40] relies on two assumptions for the true distribution with no equivalent in our article: a ρ-mixing assumption **[A⋆mixing]**, which in particular imply that the process $(X_t)_{t\geqslant 1}$ is ergodic, to obtain concentration inequalities, and a forgetting assumption **[A⋆forget]**, to truncate the dependencies in the past, that is to approximate $p_t^\star(X_t|X_1^{t-1})$ by $p_t^\star(X_t|X_{t-k}^{t-1})$ for $k < t$. These assumptions are not required for our results: we do not even need the process to be ergodic.

The other assumptions are similar in our article and [40]. The analog to the Lipschitz part of Assumption 4-HMM are Assumptions **[Aentropy]** and **[Agrowth]**, and Section B.2 follows a proof similar to the one of Lemma 6 in order to control the entropy of the class of log-likelihoods. The analogs of our assumptions that $p_t^\star(X_t)$ has a finite moment and that the g_η's are lower and upper bounded in Assumption 4-HMM are **[A⋆tail]** and **[Atail]**; **[Atail]** is actually slightly more general and could be used as is here, but we chose to keep Assumption 4-HMM simple. Finally, **[Aergodic]** is an integral part of the way the models Θ_m are defined.

If the process $(X_t)_t$ is ergodic, the risk $\mathbf{K}$ used by [40] is actually the limit of $\mathbb{E}[\mathbf{K}_n]$ when $n \to +\infty$. If in addition a forgetting assumption such as **[A⋆forget]** in [40] holds for $p^\star$, then

$$|\mathbf{K} - \mathbb{E}[\mathbf{K}_n]| = \mathcal{O}((\log n)^4/n),$$

which is negligible compared to the residual term of our oracle inequalities, so that both risks can be used interchangeably.

4.3. *Models of neuronal networks in discrete-time.* Neurons are electrical cells that communicate via the emission of action potentials, also called spikes [32]. The shape of the action potential is essentially constant and the important information is the time at which the spikes are emitted. Time is discretized, so the network is represented by a process $(X_t^i)_{t\in\mathbb{Z}, i\in I}$, where I is the set of all neurons constituting the network and $X_t^i = 1$ if the neuron i spikes at time t and $X_t^i = 0$ otherwise. We consider the filtration $\mathcal{F}_t = \sigma((X_s^i)_{s\leqslant t, i\in I})$.

Since communication between neurons is not instantaneous, most authors [32, 43] usually assume that conditionally to $\mathcal{F}_{t-1}$, the $(X_t^i)_{i\in I}$ are independent, so that the whole activity can be described by just giving the distributions $p_t^{\star,i}$ with

$$\forall i \in I,\ \forall t \in \mathbb{Z}, \quad p_t^{\star,i} = \mathbb{P}(X_t^i = 1 \,|\, \mathcal{F}_{t-1}).$$

We assume the process to be stationary.

One of the main neuronal model in discrete-time is the discrete Hawkes process [43], which can be modeled by

$$p_t^{i,H} = \phi_i \left(\sum_{j \in I} \sum_{s<t} h_{j \to i}(t-s) X_s^j \right),$$

where ϕ_i is a rate function that is usually assumed to be Lipschitz, increasing and taking values in $[0,1]$, and where $h_{j \to i}$ are interaction functions: if it is positive at delay δ, neuron j excites neuron i after a delay δ; if it is negative, neuron j inhibits neuron i after a delay δ. For instance, the linear case is the situation where $\phi_i(x) = \mu_i + x$. In the sequel, to simplify, $\phi_i(.)$ is supposed to be fixed and known. Only the functions $h_{j \to i}$ are unknown.

The Galves-Löcherbach neuronal model [32, 31] is slightly different, here

$$p_t^{i,GL} = \phi_i \left(\sum_{j \in I} \sum_{s=L_t^i}^{t-1} h_{j \to i}(t-s) X_s^j \right),$$

where L_t^i is the time of the last spike of neuron i. In contrast to the Hawkes process, the neurons of this model essentially reset their memory each time they spike.

In practice, only a small finite subset F of I is observed on a finite time duration, say $t = -A, \ldots, n$, for some positive A. For a fixed $i \in F$, we are interested by estimating $(p_t^{\star,i})_{t \in [n]}$ based on the observations of $(X_t^j)_{-A \leqslant t \leqslant n, j \in F}$.

We are interested in a specific neuron of interest $i \in F$, so the process $(X_t)_{t \in [n]}$ from our oracle inequalities is taken to be $(X_t^i)_{t=1,\ldots,n}$. The filtration $\mathcal{F}_t$ is the one defined above and generated by the whole network. Finally, in order to define the models, we have access to more information that $(X_t^i)_{t=1,\ldots,n}$ but less than the whole network: we may only use the observations $(X_t^j)_{j \in F, t=-A,\ldots,n}$, which are indeed $\mathcal{F}_t$ adapted.

Whatever the neuronal model (p^H for Hawkes or p^{GL} for Galves-Löcherbach) that we choose, we need to parameterize it. We define model m by choosing a finite subset of F, called V_m, which is the proposed neighborhood for neuron i in model m, and by choosing $A_m \leqslant A$ a maximal lag of interaction. In model m, all the $h_{j \to i}(u)$ are null if $j \notin V_m$ or $u > A_m$, so the model is parameterized by $(\theta_{j,u} := h_{j \to i}(u))_{j \in V_m, u=1,\ldots,A_m} \in \mathbb{R}^{A_m |V_m|}$. Given this parameterization, the conditional distributions are defined by

$$p_t^{i,m,H} = \phi_i \left(\sum_{j \in V_m} \sum_{u=1}^{A_m} \theta_{j,u} X_{t-u}^j \right)$$

and

$$p_t^{i,m,GL} = \phi_i \left(\sum_{j \in V_m} \sum_{u=1}^{\min(A_m, t-L_t^i)} \theta_{j,u} X_{t-u}^j \right).$$

4.3.1. *Validation of the assumptions.* Assumption 1 means that $p_t^{\star,i}$ as well as $p_t^{i,m,H}$ or $p_t^{i,m,GL}$ are in $[\varepsilon, 1-\varepsilon]$. This is a very common assumption in these settings (see [30, 43]). In this sense, Assumption 1bis can be seen as a relaxation with respect to previous works.

The assumption that ϕ_i is Lipschitz with constant L is a very classical one [30]. Together with Assumption 1, it implies that $\log(\phi_i)$ (probability of a spike) and $\log(1-\phi_i)$ (probability of no spike) are Lipschitz with constant $2\varepsilon^{-1}L$. Thus, the first part of Assumption 2 is satisfied with $L_m = 2\varepsilon^{-1}L$ and the ℓ_1 norm $\|\theta\|_1$.

For the second part of Assumption 2, it depends on ϕ_i. Indeed, since ϕ_i is increasing, we can define

$$\Theta_m = \left\{ \theta \in \mathbb{R}^{A_m|V_m|} \text{ such that } \varepsilon \leqslant \phi_i \left(\sum_{j,u} \theta_{j,u} \mathbf{1}_{\theta_{j,u}<0} \right) \text{ and } \phi_i \left(\sum_{j,u} \theta_{j,u} \mathbf{1}_{\theta_{j,u}>0} \right) \leqslant 1-\varepsilon \right\},$$

to ensure that Assumption 1 is satisfied. Since ϕ_i is increasing and Lipschitz, we can define its inverse, so that if ε and $1-\varepsilon$ are possible values for ϕ_i (as is typically the case for linear or sigmoid functions) then it automatically follows that for all $\theta \in \Theta_m$

$$\|\theta\|_1 \leqslant |\phi_i^{-1}(\varepsilon)| + |\phi_i^{-1}(1-\varepsilon)|,$$

and so the second part of Assumption 2 is satisfied with $M_m = |\phi_i^{-1}(\varepsilon)| + |\phi_i^{-1}(1-\varepsilon)|$.

Assumption 3 is fulfilled since the parametrizations are continuous. Assumption 4 is automatically fulfilled in each of the models because $p_t^{i,m,H}$ and $p_t^{i,m,GL}$ only depend on a finite set of X_s^j. Assumption 4 for $p^\star$ can be solved by assuming the following continuity assumption, which is standard in this setting (see [30, 32]). Let x be a past configuration, i.e. a possible value for $(X_s^j)_{j\in I, s<t}$, and let us remark that by stationarity, $p_t^{\star,i}$ can be seen as a function of x and not of t:

$$p^{\star,i}(x) = \mathbb{P}(X_t^i = 1 | (X_s^j)_{j\in I, s<t} = x).$$

The continuity assumption of the neuronal model assumes that there exists a nested sequence $(S_k)_{k\geqslant 1}$ of finite subsets of $I \times \mathbb{Z}_-$ such that $S_k \underset{k\to\infty}{\longrightarrow} I \times \mathbb{Z}$ and such that

$$\sup\{|p_t^{\star,i}(x) - p_t^{\star,i}(y)| \text{ such that } x_{|S_k} = y_{|S_k}\} \underset{k\to\infty}{\longrightarrow} 0,$$

where $x_{|S_k}$ is the configuration restricted to the indices in S_k. Informally, this continuity assumption states that one can approximate $p^{\star,i}(x)$ by what happens on a finite number of x_s^j. To check Assumption 4, it is sufficient to take as $\mathcal{Q}$ the countable set of x such that there exists k for which x is null outside of S_k.

4.3.2. *Result.* Following Theorem 2 or its corollary, and since A_m does not depend on m, one can take

$$\mathrm{pen}(m) = \mathcal{O}\left(\log(n)^2 \frac{D_m}{n} \right).$$

Under mild conditions (see for instance [32] or [43]), these processes are stationary and $\mathbb{E}(\mathbf{K}_n(p_\theta^m))$ does not depend on n. However we need to compare $\mathbf{K}_n(p_\theta^m)$ with the ℓ^2 distance to compare our results with existing ones. To do so, let us use the following result.

LEMMA 7. *Suppose that Assumption 1 holds for some $\varepsilon > 0$. Let $p = (p_t)_{t\in[n]}$ be any sequence of conditional densities satisfying for all $t \in [n]$, $p_t(X_t) \in [\varepsilon, \varepsilon^{-1}]$. Then, for all $t \in [n]$,*

$$\frac{\varepsilon^2}{2} \mathbb{E}\left[\log^2 \frac{p_t^\star(X_t)}{p_t(X_t)} \middle| \mathcal{F}_{t-1} \right] \leqslant \mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_t(X_t)} \middle| \mathcal{F}_{t-1} \right] \leqslant \frac{1}{2\varepsilon^2} \mathbb{E}\left[\log^2 \frac{p_t^\star(X_t)}{p_t(X_t)} \middle| \mathcal{F}_{t-1} \right].$$

Since $-\log(p_{\theta,t}^m / p_t^\star)$ is upper and lower bounded, Lemma 7 shows that

$$\mathbb{E}\left[\log^2 \frac{p_t^\star(X_t)}{p_{\theta,t}^m(X_t)} \middle| \mathcal{F}_{t-1} \right] \lesssim \mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_{\theta,t}^m(X_t)} \middle| \mathcal{F}_{t-1} \right] \lesssim \mathbb{E}\left[\log^2 \frac{p_t^\star(X_t)}{p_{\theta,t}^m(X_t)} \middle| \mathcal{F}_{t-1} \right], \tag{7}$$

where $\lesssim$ means that the inequality holds up to positive multiplicative constant.

Moreover, whatever the neuronal model (Hawkes or Galves-Löcherbach), we can expand $\theta^m \in \Theta^m$ with zeroes so that it is defined on $I \times \mathbb{N}^*$. If ϕ_i has a derivative that is upper and lower bounded by some positive constant, so do $\log(\phi_i)$ and $\log(1-\phi_i)$, and therefore, whatever the value of X_t,

$$\left|\sum_{j\in I}\sum_{s<t}(h_{j\to i}(t-s)-\theta^m_{j,t-s})X^j_s\right| \lesssim \left|\log\left(\frac{p^m_{\theta^m,t}(X_t)}{p^*_t(X_t)}\right)\right| \lesssim \left|\sum_{j\in I}\sum_{s<t}(h_{j\to i}(t-s)-\theta^m_{j,t-s})X^j_s\right|$$

(for the Hawkes case and with a restricted sum in the lower bound in the Galves-Löcherbach case). Both upper and lower bounds are $\mathcal{F}_{t-1}$ measurable. Hence, going back to the oracle inequality, we can express both the upper bound and the lower bound in terms of the average square distance

$$\frac{1}{n}\sum_{t=1}^n\left|\sum_{j\in I}\sum_{s<t}(h_{j\to i}(t-s)-\theta^m_{j,t-s})X^j_s\right|^2,$$

and the upper bound of the oracle inequality is a trade-off between the bias measured by the average square distance above and the penalty in D_m/n up to logarithmic terms.

Let us compare this result to the ones in [43] for Hawkes and Galves-Löcherbach process. In [43], the authors can only consider linear models (i.e. ϕ_i is linear) and use least-square contrast on the p. They perform variable selection thanks to an ℓ_1 penalization, whereas in our case the summability condition on Σ makes it impossible to perform variable selection by considering the full set of subsets of variables. Despite this difference, their oracle inequality is for the exact same average square distance as mentioned above with, in the upper bound, a trade-off between the bias and a term in D_m/n up to logarithmic terms. However, in their case, the dimension D_m has to be smaller than a given a priori level of sparsity s. Moreover, their constant in front of D_m is given by an RE inequality on the Gram matrix. In the most general case considered by [43], this constant depends on the size of the observed network F and explodes with the size of F; in the Hawkes case, this bound depends on s, and explodes for moderate s, whereas in our case the penalty can handle large D_m thanks to the summability condition in Σ which ensures that the number of models remains reasonable.

For the Galves-Löcherbach model, we can also compare this result with the ones of [30]. In [30] (see also [28] for recent improvements), the authors envision the estimation of the interaction neighborhood of a neuron i. In these two articles, the authors assume that the set of observed neurons F contains the interaction neighborhood of i. In other words, with their assumptions on the process, at least one model is well specified. Our result shows that we can estimate the conditional probability of spiking for i with parametric rate in D_m/n up to logarithmic terms, as the Wilks phenomenon would predict if it was applicable in this case, but without even knowing in advance which is the true model. However, it is not clear if this means that the neighborhood of interaction is correctly estimated by $V_{\hat{m}}$. We cannot manage variable selection per se because the complexity of the family of models would be too large for our general result. But we can at least hope that $V_{\hat{m}}$ would contain the true interaction neighborhood of neuron i with high probability [2]. On the other hand, our procedure is much

[2] This would probably require additional assumptions, such as minimal strength of the interaction inside the neighborhood as in [30] and [28].

less demanding on computing resources than theirs and does not require the precise examination of 0 and 1 patterns, which in practice usually prevents them from using it with more than 4 observed neurons [17]. Thus, our method could at least be used to restrict the set of observed neurons before using the methods developed in [30, 28].

4.4. *Adversarial multi-armed bandits as a cognitive learning model.* All previous applications concerned the stationary case, enabling comparisons to be made with existing results. To illustrate that the oracle inequalities of Section 3 apply even in a non-stationary framework, let us look at another configuration: learning data. Here, the goal is to estimate how an individual or system learns to perform a task by observing its training as it unfolds. The data produced cannot, in essence, be independent or stationary. Although many authors have advocated the use of MLE in practice [23, 56], this problem was studied theoretically for the first time in [5]. In this previous work, only the property of the MLE on one model is studied, without the model selection part.

The single model studied in [5] is the `Exp3` algorithm [6]. In the Machine Learning literature, this algorithm was originally proposed to solve the adversarial multi-armed bandit problem, which is a game played sequentially between a learner and an adversary (the environment) [18, 39] . At each round the learner must choose an action X_t from a set of actions $[K]$ for some integer $K \geqslant 1$ and the adversary returns a loss $g_{X_t,t}$ for this action. The term "adversarial" comes from the fact that the loss $g_{X_t,t}$ incurred in round t may depend on the player's past actions.

Algorithm 1 `Exp3` with learning rate η

Initialization: $p_{\eta,1} = \left(\frac{1}{K}, \dots, \frac{1}{K}\right)$.
for $t = 1, 2, \dots$ **do**
  Draw an action $X_t \sim p_{\eta,t}$ and receive a loss $g_{X_t,t} \in [0, 1]$.
  Update for all $k \in [K]$:

$$p_{\eta,t+1}(k) = \frac{\exp\left(-\eta \sum_{s=1}^{t} \hat{L}_{\eta,s}(k)\right)}{\sum_{j=1}^{K} \exp\left(-\eta \sum_{s=1}^{t} \hat{L}_{\eta,s}(j)\right)} \qquad \text{where } \hat{L}_{\eta,s}(j) = \frac{g_{j,s}}{p_{\eta,s}(j)} \mathbf{1}_{X_s=j}$$

The fact that the learner plays as if the environment is adversarial makes it a realistic model for cognitive processes where humans or animals need to adapt to a changing environment. However, in a cognitive experiment, most of the time the environment is not adversarial, even if the learner does not know it and uses an adversarial strategy anyway.

4.4.1. *Comparison with recent work.* In [5], we consider situations where $g_{k,t} = g_k$ depends only on k for all $k \in [K]$, as is often the case for cognitive experiments. We show that if the learning rate $\eta = \theta$ is fixed, no estimator can achieve polynomial rates of convergence. This is mainly due to the fact that $p_{\eta,t}(k)$ can rapidly become absurdly small, and as a consequence, only an extremely small number of X_t are relevant for estimation. If an individual has finished learning and is no longer making mistakes, it is impossible to improve the estimate of his learning behavior, even by continuing to observe him. Therefore, in [5], another asymptotic in T is proposed. The goal is to estimate a new parameter $\theta \in [r, R]$ such that the learning rate η in `Exp3` is equal to $\eta = \theta/\sqrt{T}$. The estimation is based on the first

$$T_\varepsilon = \left\lfloor \left(\frac{1}{K} - \varepsilon\right) \frac{\sqrt{T}}{R} \right\rfloor$$

observations because up to T_ε, all the $p_{\frac{\theta}{\sqrt{T}},t}(k)$ are greater than ε, a fixed constant in $(0,1)$ meant to prevent the various probabilities from vanishing. The choice of the rate $\sqrt{T}$ can be relaxed by other powers of T, but for the purpose of this illustration, we consider only this case, which also corresponds to the renormalization of the learning rate for which `Exp3` satisfies sublinear regret bounds [18].

With this renormalization, we proved in [5] a convergence in $T_\varepsilon^{-1/2}$ for

$$\frac{1}{T_\varepsilon}\sum_{t=1}^{T_\varepsilon}\sum_{k=1}^{K}(p_{\frac{\theta}{\sqrt{T}},t}(k) - p_{\frac{\hat\theta}{\sqrt{T}},t}(k))^2,$$

with θ the true parameter and $\hat\theta$ the MLE. Let us compare this with Theorem 2 or its corollary in the case of a single, well specified model, where $p_t^\star = p_{\frac{\theta}{\sqrt{T}},t}$ for some $\theta \in [r,R]$. Take $\mathcal{F}_t = \sigma(X_1^t)$ and $n = T_\varepsilon$, so that Assumption 1 is directly satisfied. Since $\log$ is Lipschitz on $[\varepsilon,1]$, Assumption 2 follows as a direct consequence of Lemma 4.3 of [5], with respect to $\|\theta\|_\infty$, with a constant L that only depends on R and ε. Moreover, we can always take $M = R$. Finally, Assumption 4 is trivial since there is only a finite set of values for X_t.

Along with Lemma 8, Theorem 2 shows that $\mathbf{K}_n(\tilde p)$ is of order $\log(T_\varepsilon)^2/T_\varepsilon$, which is faster than [5]. Indeed, Lemma 8 shows that $\mathbf{K}_n(p)$ is comparable to the square norm used in [5].

LEMMA 8. *Let $\varepsilon \in (0,1)$. Assume that for all $t \in [T_\varepsilon]$ and $k \in [K]$, $p_t^\star(k) \geqslant \varepsilon$. Let $p = (p_t)_{t\in[n]}$ be a sequence of conditional distributions over $[K]$ satisfying for all $t \in [T_\varepsilon]$ and all $k \in [K]$, $p_t(k) \geqslant \varepsilon$. Then,*

$$\mathbf{K}_n(p) \geqslant \frac{1}{8T_\varepsilon}\sum_{t=1}^{T_\varepsilon}\sum_{k=1}^{K}\left(p_t^\star(k) - p_t(k)\right)^2.$$

4.4.2. *A model selection framework.* Theorem 2 or its corollary go further than [5] by also allowing for model selection and bias. An important application in a cognitive learning experience is to estimate the reward and to identify the granularity or precision the learner is capable of achieving in his actions. Indeed, it is rarely obvious to assign a precise numerical value to a real-life reward, so that it really reflects the impact it has on the learner. For example, what is the value of crabmeat to an octopus or chocolate to a rodent? Nor is it obvious how precise the learner is when choosing an action, whether the number of possible actions is continuous or very high. Since it is impossible to reproduce a given action perfectly, almost identical actions should provide almost identical rewards; how different actions have to be to be considered distinct?

One possibility is to use `Exp3` with a different parameterization. For a given model m, consider a partition $\mathcal{I}^m$ of the set of possible actions (which may even be continuous). Let D_m be the number of disjoint sets in $\mathcal{I}^m$. The parameters of model m are the rewards of each possible set I of the partition, that we call g_I^m. Since the role of the learning parameter η introduced in Algorithm 1 is redundant with that of the (unknown) rewards g, we remove it from the model. Thus, under model m, the learner proceeds as follows:

- Initialize $p_1^m = (1/D_m, \dots, 1/D_m)$
- for $t \geqslant 1$,
  - pick the set I_t according to p_t^m and pick X_t inside I_t uniformly
  - update for all $I \in \mathcal{I}^m$

$$p_{t+1}^m(I) = \frac{\exp\left(-\sum_{s=1}^t \hat L_s^m(I)\right)}{\sum_{J=1}^{D_m}\exp\left(-\sum_{s=1}^t \hat L_s^m(J)\right)} \qquad \text{where} \qquad \hat L_s^m(J) = \frac{g_J^m}{p_s(J)}\mathbf{1}_{X_s\in J}. \tag{8}$$

First, it is important to note that for such model, $\hat{L}_t^m(J)$ is still an unbiased estimator of the loss g_J^m with respect to the distribution p_s, as it is the case for Exp3-type bandit algorithms (see Section F for more details). Thus, these models make sense from a pure adversarial bandit point of view.

To obtain a meaningful renormalization as before, we call $p_{\theta^m,t}^m$ the distribution of X_t given $\mathcal{F}_{t-1} = \sigma(X_1^{t-1})$ when $g_J^m = \frac{\theta_J^m}{\sqrt{T}}$ and let $\hat{L}_{\theta^m,s}^m(J)$ be the expected loss instead of $\hat{L}_s^m(J)$.

4.4.2.1. Validation of the assumptions. The following proposition shows that Assumption 1 is satisfied provided that the number of sets D_m in the partition $\mathcal{I}^m$ is uniformly bounded for all $m \in \mathcal{M}$. A proof of this result is available in Section F.2.2.

PROPOSITION 9. *Assume that $D := \sup_{m\in\mathcal{M}} D_m < +\infty$. Then, Assumption 1 is satisfied for any $\varepsilon \in (0, 1/D)$ as long as $n = T_\varepsilon$ where*

$$T_\varepsilon = \left\lfloor \left(\frac{1}{D} - \varepsilon\right) \frac{\sqrt{T}}{R} \right\rfloor.$$

The following proposition is an adaptation of Lemma 4.3 of [5] to show that Assumption 2 is satisfied as long as all the θ_J^m are in $[r, R]$. Its proof is available in Section F.2.3.

PROPOSITION 10. *Let D and T_ε be the constants of Proposition 9. Assumption 2 is satisfied for the norm $\|\cdot\|_\infty$ on $\mathbb{R}^{D_m}$ for any $\varepsilon \in (0, 1/D)$ as long as $n = T_\varepsilon$, $M_m = M = R$ and*

$$L_m = L = \frac{1}{D\varepsilon^2} \exp\left(\frac{1}{D\varepsilon^2}\right).$$

Assumptions 3 and 4 are again easily satisfied even if the set of actions is continuous.

4.4.2.2. Result. One can choose a penalty of order

$$\mathrm{pen}(m) = \mathcal{O}\left(\log(n)^2 \frac{D_m}{n}\right),$$

and obtain an oracle inequality on $\mathbf{K}_n(\tilde{p})$, which as seen above implies an oracle inequality on the square distance. In the end, we obtain a way of estimating at the same time

- with the partition $\mathcal{I}_{\hat{m}}$: the precision in the perception and execution of actions, i.e. which sets of actions are confused and which are considered distinct, and how precise the learner is able to be when choosing actions,
- with the resulting $\hat{\theta}_J^{\hat{m}}/\sqrt{T_\varepsilon}$: the estimated reward, i.e. the numerical value that quantifies the average impact of the action result on the learner's behavior, modeled by a piecewise constant function on the partition $\mathcal{I}_{\hat{m}}$.

Up to our knowledge, this is the first model selection result of this type for learning data.

5. Conclusion. In summary, our work presents some oracle inequalities based on a penalized log-likelihood criterion under weak assumptions, which are effective even for dependent, non-ergodic and non-stationary processes. Although these assumptions may seem restrictive, they allow us to draw significant conclusions. In fact, these assumptions are often comparable to the regularity conditions generally assumed to obtain consistency and asymptotic normality in MLE, such as the Hölder conditions.

Our results have been applied in a variety of contexts, including i.i.d. processes, hidden Markov models (HMMs), neural networks and reinforcement learning (RL). However, our approach has certain limitations. Firstly, the penalty is only known up to a constant, which is a common problem, but there are data-driven heuristics to overcome this. Secondly, our stochastic risk $\mathbf{K}_n$ is generally difficult to understand intuitively. Nevertheless, in the simplest cases, it is related to more familiar criteria such as classical Kullback-Leibler divergence or more traditional ℓ^2 distance (as shown in Lemma 8).

In the future, it would be interesting to study similar results using other criteria, such as empirical risk minimization, instead of focusing solely on likelihood maximization. This would help us to understand the extent to which we are relying on the specific properties of log-likelihood, and to consider extensions to other types of risk.

Acknowledgments. This work was supported by the French government, through the UCA$^{\text{Jedi}}$ and 3IA Côte d'Azur Investissements d'Avenir managed by the National Research Agency (ANR-15- IDEX-01 and ANR-19-P3IA-0002), by the interdisciplinary Institute for Modeling in Neuroscience and Cognition (NeuroMod) of the Université Côte d'Azur and directly by the National Research Agency ANR-19-CE40-0024 with the ChaMaNe project.

REFERENCES

[1] AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. *2nd international symposium on information theory, Akademia Kiado, Budapest* 267–281.

[2] ARLOT, S. (2019). Minimal penalties and the slope heuristics: a survey. *J. SFdS* **160** 158–168. MR4021422

[3] ARLOT, S. and LERASLE, M. (2016). Choice of V for V-fold cross-validation in least-squares density estimation. *J. Mach. Learn. Res.* **17** Paper No. 208, 50. MR3595142

[4] AUBERT, J. and LEHÉRICY, L. (2024). Exponential inequalities for suprema of processes with stochastic normalization. preprint.

[5] AUBERT, J., LEHÉRICY, L. and REYNAUD-BOURET, P. (2023). On the convergence of the MLE as an estimator of the learning rate in the Exp3 algorithm. In *Proceedings of the 40th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* **202** 1244–1275. PMLR.

[6] AUER, P., CESA-BIANCHI, N., FREUND, Y. and SCHAPIRE, R. E. (2002). The nonstochastic multiarmed bandit problem. *SIAM journal on computing* **32** 48–77.

[7] BARAUD, Y. (2010). A Bernstein-type inequality for suprema of random processes with applications to model selection in non-Gaussian regression. *Bernoulli* **16** 1064 – 1085.

[8] BARAUD, Y. and BIRGÉ, L. (2016). Rho-estimators for shape restricted density estimation. *Stochastic Process. Appl.* **126** 3888–3912. https://doi.org/10.1016/j.spa.2016.04.013 MR3565484

[9] BAUM, L. E. and PETRIE, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The annals of mathematical statistics* **37** 1554–1563.

[10] BERCU, B., DELYON, B., RIO, E. et al. (2015). *Concentration inequalities for sums and martingales*. Springer.

[11] BERCU, B. and TOUATI, A. (2008). Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* **18** 1848–1869. https://doi.org/10.1214/07-AAP506 MR2462551

[12] BERCU, B. and TOUATI, T. (2019). New insights on concentration inequalities for self-normalized martingales. *Electronic Communications in Probability* **24** 1 – 12.

[13] BICKEL, P. J. and DOKSUM, K. A. (2015). *Mathematical statistics—basic ideas and selected topics. Vol. 1*, second ed. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3445928

[14] BIRGÉ, L. (1983). Approximation dans les espaces métriques et théorie de l'estimation. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **65** 181–237.

[15] BIRGÉ, L. and MASSART, P. (2001). Gaussian model selection. *J. Eur. Math. Soc. (JEMS)* **3** 203–268. https://doi.org/10.1007/s100970100031 MR1848946

[16] BOUCHERON, S. and MASSART, P. (2011). A high-dimensional Wilks phenomenon. *Probab. Theory Related Fields* **150** 405–433. https://doi.org/10.1007/s00440-010-0278-7 MR2824862

[17] BROCHINI, L., GALVES, A., HODARA, P., OST, G. and POUZAT, C. (2016). Estimation of neuronal interaction graph from spike train data. *arXiv preprint arXiv:1612.05226*.

[18] BUBECK, S. and CESA-BIANCHI, N. (2012). Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. *CoRR* **abs/1204.5721**.

[19] BURNHAM, K. P. and ANDERSON, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33** 261–304. https://doi.org/10.1177/0049124104268644 MR2086350

[20] CASTELLAN, G. (2003). Density estimation via exponential model selection. *IEEE Trans. Inform. Theory* **49** 2052–2060. https://doi.org/10.1109/TIT.2003.814485 MR2004713

[21] CLARK, D. S. (1987). Short proof of a discrete gronwall inequality. *Discrete Applied Mathematics* **16** 279-281. https://doi.org/10.1016/0166-218X(87)90064-3

[22] COX, D. R. (1975). Partial likelihood. *Biometrika* **62** 269–276.

[23] DAW, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision making, affect, and learning: Attention and performance XXIII* **23**.

[24] DE CASTRO, Y., GASSIAT, É. and LACOUR, C. (2016). Minimax Adaptive Estimation of Nonparametric Hidden Markov Models. *Journal of Machine Learning Research* **17** 1–43.

[25] DE CASTRO, Y., GASSIAT, É. and LE CORFF, S. (2017). Consistent estimation of the filtering and marginal smoothing distributions in nonparametric hidden Markov models. *IEEE Transactions on Information Theory*.

[26] DE LA PEÑA, V. H., KLASS, M. J. and LEUNG LAI, T. (2004). Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *The Annals of Probability* **32**.

[27] DE LA PEÑA, V. H., LAI, T. L. and SHAO, Q.-M. (2009). *Self-normalized processes*. *Probability and its Applications (New York)*. Springer-Verlag, Berlin Limit theory and statistical applications. https://doi.org/10.1007/978-3-540-85636-8 MR2488094

[28] DE SANTIS, E., GALVES, A., NAPPO, G. and PICCIONI, M. (2022). Estimating the interaction graph of stochastic neuronal dynamics by observing only pairs of neurons. *Stochastic Processes and their Applications* **149** 224–247.

[29] DOUC, R., MOULINES, É. and RYDÉN, T. (2004). Asymptotic Properties of the Maximum Likelihood Estimator in Autoregressive Models with Markov Regime. *Annals of Statistics* 2254–2304.

[30] DUARTE, A., GALVES, A., LÖCHERBACH, E. and OST, G. (2019). Estimating the interaction graph of stochastic neural dynamics. *Bernoulli* **25** 771–792.

[31] GALVES, A. and LÖCHERBACH, E. (2013). Infinite Systems of Interacting Chains with Memory of Variable Length-A Stochastic Model for Biological Neural Nets. *Journal of Statistical Physics* **151**.

[32] GALVES, A., LÖCHERBACH, E. and POUZAT, C. (2024). *Probabilistic Spiking Neuronal Nets – Data, Models and Theorems*. Springer.

[33] GAO, B. and PAVEL, L. (2017). On the properties of the softmax function with application in game theory and reinforcement learning. *arXiv preprint arXiv:1704.00805*.

[34] GHOSAL, S. and VAN DER VAART, A. (2007). Posterior convergence rates of Dirichlet mixtures at smooth densities. *Ann. Statist.* **35** 697–723.

[35] HANSEN, N. R., REYNAUD-BOURET, P. and RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21**. https://doi.org/10.3150/13-bej562

[36] HOUDRÉ, C. and REYNAUD-BOURET, P. (2002). Exponential Inequalities for U-Statistics of Order Two with Constants. *Stochastic Inequalities and Applications. Progress in Probability* **56**.

[37] LACOUR, C. (2008). Nonparametric estimation of the stationary density and the transition density of a Markov chain. *Stochastic Process. Appl.* **118** 232–260. https://doi.org/10.1016/j.spa.2007.04.013 MR2376901

[38] LACOUR, C. (2008). Adaptive estimation of the transition density of a particular hidden Markov chain. *J. Multivariate Anal.* **99** 787–814. https://doi.org/10.1016/j.jmva.2007.04.006 MR2405092

[39] LATTIMORE, T. and SZEPESVÁRI, C. (2020). *Bandit Algorithms*. Cambridge University Press.

[40] LEHÉRICY, L. (2021). Nonasymptotic control of the MLE for misspecified nonparametric hidden Markov models. *Electronic Journal of Statistics* **15** 4916–4965.

[41] LEHMANN, E. L. (1999). *Elements of large-sample theory*. *Springer Texts in Statistics*. Springer-Verlag, New York. https://doi.org/10.1007/b98855 MR1663158

[42] MASSART, P. (2007). *Concentration inequalities and model selection: Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003*. Springer.

[43] OST, G. and REYNAUD-BOURET, P. (2020). Sparse space-time models: concentration inequalities and Lasso. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2377–2405. https://doi.org/10.1214/19-AIHP1042 MR4164841

[44] RABINER, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77** 257–286.

[45] RAMPONI, G., DRAPPO, G. and RESTELLI, M. (2020). Inverse reinforcement learning from a gradient-based learner. *Advances in Neural Information Processing Systems* **33** 2458–2468.

[46] SHEN, W., TOKDAR, S. T. and GHOSAL, S. (2013). Adaptive Bayesian multivariate density estimation with Dirichlet mixtures. *Biometrika* **100** 623–640.

[47] SPOKOINY, V. (2012). Parametric estimation. Finite sample theory. *Ann. Statist.* **40** 2877–2909. https://doi.org/10.1214/12-AOS1054 MR3097963

[48] SPOKOINY, V. (2017). Penalized maximum likelihood estimation and effective dimension. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 389–429. https://doi.org/10.1214/15-AIHP720 MR3606746

[49] SUR, P. and CANDÈS, E. J. (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proc. Natl. Acad. Sci. USA* **116** 14516–14525. https://doi.org/10.1073/pnas.1810420116 MR3984492

[50] SUR, P., CHEN, Y. and CANDÈS, E. J. (2019). The likelihood ratio test in high-dimensional logistic regression is asymptotically a *rescaled* chi-square. *Probab. Theory Related Fields* **175** 487–558. https://doi.org/10.1007/s00440-018-00896-9 MR4009715

[51] TALAGRAND, M. (1996). New concentration inequalities in product spaces. *Invent. Math.* **126** 505–563. https://doi.org/10.1007/s002220050108 MR1419006

[52] TANG, C. Y. and LENG, C. (2010). Penalized high-dimensional empirical likelihood. *Biometrika* **97** 905–919. With supplementary material available online. https://doi.org/10.1093/biomet/asq057 MR2746160

[53] VAN DE GEER, S. (1995). Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *Ann. Statist.* **23** 1779–1801. https://doi.org/10.1214/aos/1176324323 MR1370307

[54] VAN DER VAART, A. W. and WELLNER, J. A. (1996). M-estimators. *Weak Convergence and Empirical Processes: With Applications to Statistics* 284–308.

[55] WILKS, S. (1938). The large-sample distribution of the likelihood ratio for composite testing hypotheses. *Ann. Math. Stat.* 60-62.

[56] WILSON, R. C. and COLLINS, A. G. E. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife* **8** e49547.

[57] ZHENG, C., FERRARI, D. and YANG, Y. (2019). Model selection confidence sets by likelihood ratio testing. *Statist. Sinica* **29** 827–851. MR3931390

APPENDIX A: PROOF OF THEOREMS 2 AND 4

In the sequel to visually keep track of important dependencies, we denote, for any $t \in [n]$, $m \in \mathcal{M}$, $\theta \in \Theta_m$, $x_t \mapsto p^m_{\theta,t}(x_t|\mathcal{F}_{t-1})$ the density of X_t conditionally to $\mathcal{F}_{t-1}$ under the parameter θ, and likewise for $p^\star_t(x_t|\mathcal{F}_{t-1})$, instead of $p^m_{\theta,t}(x_t)$ and $p^\star_t(x_t)$ respectively.

The proofs of both Theorems are nearly identical. Let us introduce the following notations in order to show both at the same time in a unified framework: for all $m \in \mathcal{M}$,

- under Assumptions 1 and 2, let $F^\infty_m = 2\log(\varepsilon^{-1})$, $F^{\text{höl}}_m = L_m$ and $q_n = 0$,
- under Assumptions 1bis and 2bis, let $F^\infty_m = (1+\alpha)B_m \log n$, $F^{\text{höl}}_m = L_m \log n$ and $q_n = 2n^{-\alpha}$,

and for all $t \in [n]$, let

$$f^t_A : x \in \mathcal{X} \mapsto \sup_{m \in \mathcal{M}} \max\left(\sup_{\delta,\theta \in \Theta_m \cup \{\star\}} \frac{\left|\log \frac{p^m_{\theta,t}(x|\mathcal{F}_{t-1})}{p^m_{\delta,t}(x|\mathcal{F}_{t-1})}\right|}{F^\infty_m}, \sup_{\substack{\delta,\theta \in \Theta_m \\ \delta \neq \theta}} \frac{\left|\log \frac{p^m_{\theta,t}(x|\mathcal{F}_{s-1})}{p^m_{\delta,t}(x|\mathcal{F}_{t-1})}\right|}{F^{\text{höl}}_m \|\delta - \theta\|^{\beta_m}_m} \right).$$

Note that since the models $\Theta_{m''}$ are parametric, they are separable. Therefore, by Assumption 3, the supremum over δ, θ in the definition of f^t_A can be replaced by a supremum of a countable family of measurable functions, thus ensuring that f^t_A is measurable since $\mathcal{M}$ is countable by definition.

Under either set of assumptions, almost surely, for all $t \in [n]$,

$$\mathbb{P}(f^t_A(X_t) < 1 \,|\, \mathcal{F}_{t-1}) \geqslant 1 - \frac{q_n}{n}. \tag{9}$$

Moreover, let A be the event $\{f^t_A(X_t) < 1 \text{ for all } t \in [n]\}$.

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcap_{t \in [n]} \{f^t_A(X_t) < 1\}\right)$$

$$
\begin{aligned}
&= 1 - \mathbb{P}\left(\bigcup_{t \in [n]} \{ f_A^t(X_t) \geqslant 1 \} \right) \\
&\geqslant 1 - \sum_{t \in [n]} \mathbb{P}\left(f_A^t(X_t) \geqslant 1 \right) \\
&\geqslant 1 - \sum_{t \in [n]} \mathbb{E}\left[\mathbb{P}\left(f_A^t(X_t) \geqslant 1 | \mathcal{F}_{t-1} \right) \right] \\
&\geqslant 1 - q_n,
\end{aligned}
$$

where

- the first inequality holds by the subadditivity property of a probability measure,
- the second inequality holds by the law of total expectation,
- the third inequality holds by (9).

In what follows, fix $m \in \mathcal{M}$ and $\bar{\theta}^m \in \Theta_m$. Consider the following functions, defined for all $t \in [n]$, $x_1^t \in \mathcal{X}^t$, $m' \in \mathcal{M}$ and $\delta \in \Theta_{m'}$ by

$$
g_{\delta,t}^{m'}(x_t, \mathcal{F}_{t-1}) = -\log\left(\frac{p_{\delta,t}^{m'}(x_t | \mathcal{F}_{t-1})}{p_t^\star(x_t | \mathcal{F}_{t-1})} \right) \mathbf{1}_{f_A^t(x_t) < 1}
$$

and write $g_\delta^{m'} = (g_{\delta,t}^{m'})_{t \in [n]}$.

For all $m' \in \mathcal{M}$, let $\hat{\theta}^{m'}$ be a maximizer of $\theta \in \Theta_{m'} \mapsto \frac{1}{n}\ell_n(\theta)$, let $\mathrm{crit}(m')$ be

$$
\mathrm{crit}(m') = -\frac{1}{n}\ell_n(\hat{\theta}^{m'}) + \mathrm{pen}(m'),
$$

and define the set $\mathcal{M}'$ as

$$
\mathcal{M}' = \{ m' \in \mathcal{M}, \quad \mathrm{crit}(m') \leqslant \mathrm{crit}(m) \}.
$$

For any family $h = (h_t)_{t \in [n]}$ of functions of X_t that may depend on the past, that is $h_t(X_t, \mathcal{F}_{t-1})$, we write

$$
\begin{cases}
P(h) = \dfrac{1}{n} \displaystyle\sum_{t=1}^n h_t(X_t, \mathcal{F}_{t-1}), \\
C(h) = \dfrac{1}{n} \displaystyle\sum_{t=1}^n \mathbb{E}[h_t(X_t, \mathcal{F}_{t-1}) \,|\, \mathcal{F}_{t-1}] \quad \text{the compensator of } P(h), \\
\nu(h) = P(h) - C(h) = \dfrac{1}{n} \displaystyle\sum_{t=1}^n \left(h_t(X_t, \mathcal{F}_{t-1}) - \mathbb{E}[h_t(X_t, \mathcal{F}_{t-1}) \,|\, \mathcal{F}_{t-1}] \right).
\end{cases}
$$

On the event A, it holds that for all $t \in [n]$, $x_1^t \in \mathcal{X}^t$, $m' \in \mathcal{M}$ and $\delta \in \Theta_{m'}$,

$$
g_{\delta,t}^{m'}(x_t, \mathcal{F}_{t-1}) = -\log\left(\frac{p_{\delta,t}^{m'}(x_t | \mathcal{F}_{t-1})}{p_t^\star(x_t | \mathcal{F}_{t-1})} \right).
$$

Therefore, by definition of crit, on the event A, for every $m' \in \mathcal{M}'$,

$$
P(g_{\hat{\theta}^{m'}}^{m'}) + \mathrm{pen}(m') \leqslant P(g_{\hat{\theta}^m}^m) + \mathrm{pen}(m) \leqslant P(g_{\bar{\theta}^m}^m) + \mathrm{pen}(m).
$$

Since $P = C + \nu$, on the event A, for every $m' \in \mathcal{M}'$,

$$
C\left(g_{\hat{\theta}^{m'}}^{m'}\right) + \nu(g_{\hat{\theta}^{m'}}^{m'}) \leqslant C\left(g_{\bar{\theta}^m}^m\right) + \mathrm{pen}(m) + \nu(g_{\bar{\theta}^m}^m) - \mathrm{pen}(m').
$$

Define the stochastic risk

$$\widetilde{\mathbf{K}}_n(p) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_t(X_t)} \mathbf{1}_{f_A^t(X_t)<1} \middle| \mathcal{F}_{t-1}\right].$$

Then, plugging the definition of $g^m_{\bar\theta^m}$ in the above leads to, on the event A, for every $m' \in \mathcal{M}'$,

$$\widetilde{\mathbf{K}}_n(p^{m'}_{\hat\theta^{m'}}) \leqslant \widetilde{\mathbf{K}}_n(p^m_{\bar\theta^m}) + \mathrm{pen}(m) - \nu(g^{m'}_{\hat\theta^{m'}}) - \mathrm{pen}(m') + \nu(g^m_{\bar\theta^m}).$$

Note that $\widetilde{\mathbf{K}}_n$ is not quite the stochastic risk $\mathbf{K}_n$ used in the Theorems, which is defined by

$$\mathbf{K}_n(p) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[\log \frac{p_t^\star(X_t)}{p_t(X_t)} \middle| \mathcal{F}_{t-1}\right].$$

Under Assumptions 1 and 2, almost surely, $\widetilde{\mathbf{K}}_n(p^{m'}_\theta) = \mathbf{K}_n(p^{m'}_\theta)$ for all $m' \in \mathcal{M}$ and $\theta \in \Theta_{m'}$. On the other hand, under Assumptions 1bis and 2bis, the following Lemma ensures a control on the difference between the two riks.

LEMMA 11. *Under Assumptions 1bis and 2bis, for $\alpha \geqslant 1$ and $n \geqslant 3$, almost surely, for any $m' \in \mathcal{M}$ and $\theta \in \Theta_{m'}$*

$$|\mathbf{K}_n(p^{m'}_\theta) - \widetilde{\mathbf{K}}_n(p^{m'}_\theta)| \leqslant 6B_{m'}(1+\alpha)\frac{\log n}{n^{1+\alpha}} = \frac{6F^\infty_{m'}}{n^{1+\alpha}}.$$

PROOF. See Section C.1 □

Therefore, on the event A,

$$\mathbf{K}_n(p^{m'}_{\hat\theta^{m'}}) \leqslant \mathbf{K}_n(p^m_{\bar\theta^m}) + \mathrm{pen}(m) - \nu(g^{m'}_{\hat\theta^{m'}}) - \mathrm{pen}(m') + \nu(g^m_{\bar\theta^m}) + \mathrm{Res}(m) + \mathrm{Res}(m') \tag{10}$$

with $\mathrm{Res}(m') = 0$ under Assumptions 1 and 2 and $\mathrm{Res}(m') = 6F^\infty_{m'} n^{-(1+\alpha)}$ under Assumptions 1bis and 2bis.

So far everything is similar to [42]. The goal is now to control $-\nu(g^{m'}_{\hat\theta^{m'}})$ and $\nu(g^m_{\bar\theta^m})$ in (10).

For any $m' \in \mathcal{M}$ and $\delta \in \Theta_{m'}$, let

$$\mathbf{V}_n(p^{m'}_\delta) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[\left(\log \frac{p_t^\star(X_t|\mathcal{F}_{t-1})}{p^{m'}_{\delta,t}(X_t|\mathcal{F}_{t-1})}\right)^2 \mathbf{1}_{f_A^t(X_t)<1} \middle| \mathcal{F}_{t-1}\right],$$

$$\mathring{A}_{m'} = F^{\text{höl}}_{m'} M^{\beta_{m'}}_{m'} + F^\infty_{m'} \quad \text{and} \quad v_{m'} = \mathring{A}_{m'}\sqrt{2n}. \tag{11}$$

and let $\sigma_{m'}$ be the solution of the equation

$$\sigma = \left(1 \wedge \frac{v_{m'}}{\sigma}\right)\sqrt{D_{m'}\left(1 + \frac{1}{\beta_{m'}}\log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right)} + \frac{\mathring{A}_{m'}}{\sigma} D_{m'}\left(1 + \frac{1}{\beta_{m'}}\log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right). \tag{12}$$

LEMMA 12. *Assume Assumption 4 holds, as well as either Assumptions 1 and 2 or Assumptions 1bis and 2bis. For any family* $(\eta_{m'})_{m'\in\mathcal{M}}$ *taking values in* $(0,1)$*, letting* $y_{m'} = \eta_{m'}^{-1}\sqrt{\sigma_{m'}^2 + x + D_{m'}}$ *for each* $m' \in \mathcal{M}$*, it holds with probability at least* $1 - 6e^{-x}\sum_{m'\in\mathcal{M}} \log(v_{m'})e^{-D_{m'}}$ *that for all* $m' \in \mathcal{M}$*,*

$$\sup_{\delta\in\Theta_{m'}} \left(\frac{|\nu(g_\delta^{m'})|}{2\mathbf{V}_n(p_\delta^{m'}) + \frac{1}{n}y_{m'}^2} \right) \leqslant 80(2\eta_{m'} + \eta_{m'}^2 \mathring{A}_{m'}). \tag{13}$$

PROOF. See Section C.2. □

LEMMA 13. *Under either Assumptions 1 and 2 or Assumptions 1bis and 2bis, almost surely, for all* $m' \in \mathcal{M}$ *and* $\delta \in \Theta_{m'}$*,*

$$\mathbf{V}_n(p_\delta^{m'}) \leqslant 16(F_{m'}^\infty)^2 \mathbf{K}_n(p_\delta^{m'}).$$

PROOF. See Section C.3. □

Fix some sequence $(\eta_{m'})_{m'\in\mathcal{M}}$ in $(0,1)$ to be determined later and for all $m' \in \mathcal{M}$ let $y_{m'} = \eta_{m'}^{-1}\sqrt{\sigma_{m'}^2 + x + D_{m'}}$. By definition of $\mathring{A}_{m'}$, it holds that

$$\log(v_{m'}) = \log(\mathring{A}_{m'}) + \frac{1}{2}\log(2n) = \begin{cases} \log(A_{m'}) + \frac{\log(2n)}{2} & \text{in Theorem 2,} \\ \log(A_{m'}) + \log(\log n) + \frac{\log(2n)}{2} & \text{in Theorem 4.} \end{cases}$$

In any case, since $n \geqslant 3$ and $A_{m'} \geqslant 2$,

$$\log(v_{m'}) \leqslant 4\log(n)\log(A_{m'}).$$

By Lemma 12 and 13, using the definition of Σ, for $x \geqslant 0$, it holds with probability at least $1 - 24\log(n)\Sigma e^{-x}$, for all $m' \in \mathcal{M}'$,

$$-\nu(g_{\hat\theta^{m'}}^{m'}) \leqslant 80(2\eta_{m'} + \eta_{m'}^2\mathring{A}_{m'}) \left(32(F_{m'}^\infty)^2 \mathbf{K}_n(p_{\hat\theta^{m'}}^{m'}) + \frac{1}{n}y_{m'}^2 \right)$$

and

$$\nu(g_{\hat\theta^m}^m) \leqslant 80(2\eta_m + \eta_m^2\mathring{A}_m) \left(32(F_m^\infty)^2 \mathbf{K}_n(p_{\hat\theta^m}^m) + \frac{1}{n}y_m^2 \right).$$

Injecting this result in (10), since the event A is of probability at least $1 - q_n$, it holds with probability at least $1 - q_n - 24\log(n)\Sigma e^{-x}$ that for all $m' \in \mathcal{M}'$,

$$\begin{aligned} &\left(1 - 2560(F_{m'}^\infty)^2(2\eta_{m'} + \eta_{m'}^2\mathring{A}_{m'})\right)\mathbf{K}_n(p_{\hat\theta^{m'}}^{m'}) \\ &\quad\leqslant \left(1 + 2560(F_m^\infty)^2(2\eta_m + \eta_m^2\mathring{A}_m)\right)\mathbf{K}_n(p_{\hat\theta^m}^m) + \mathrm{Res}(m) + \mathrm{Res}(m') \\ &\qquad + 80\left(\frac{2}{\eta_m} + \mathring{A}_m\right)\frac{1}{n}(\sigma_m^2 + x + D_m) + \mathrm{pen}(m) \\ &\qquad + 80\left(\frac{2}{\eta_{m'}} + \mathring{A}_{m'}\right)\frac{1}{n}(\sigma_{m'}^2 + x + D_{m'}) - \mathrm{pen}(m'). \end{aligned} \tag{14}$$

Let $\kappa \in (0,1]$ and for all $m' \in \mathcal{M}$, let

$$\eta_{m'} = \frac{c\kappa}{\mathring{A}_{m'}F_{m'}^\infty} \tag{15}$$

for some numerical constant $c > 0$ to be determined. Let us focus on the first multiplicative term on the left hand side of (14). With that choice of $\eta_{m'}$,

$$\begin{aligned} 2560(F^{\infty}_{m'})^2(2\eta_{m'} + \eta^2_{m'}\mathring{A}_{m'}) &= 2560(F^{\infty}_{m'})^2\left(\frac{2c\kappa}{\mathring{A}_{m'}F^{\infty}_{m'}} + \frac{c^2\kappa^2}{\mathring{A}^2_{m'}(F^{\infty}_{m'})^2}\mathring{A}_{m'}\right) \\ &= 2560\left(\frac{2c\kappa F^{\infty}_{m'}}{\mathring{A}_{m'}} + \frac{c^2\kappa^2}{\mathring{A}_{m'}}\right) \\ &\leqslant 2560\left(2c\kappa + c^2\kappa^2\right) \quad \text{since } F^{\infty}_{m'} \leqslant \mathring{A}_{m'} \quad \text{and} \quad \mathring{A}_{m'} \geqslant 2 \\ &\leqslant 2560\left(2c + c^2\right)\kappa \quad \text{since } \kappa \in (0,1]. \end{aligned}$$

Pick c small enough so that $2560\left(2c + c^2\right) \leqslant 1$. Therefore,

$$2560(F^{\infty}_{m'})^2(2\eta_{m'} + \eta^2_{m'}\mathring{A}_{m'}) \leqslant \kappa. \tag{16}$$

Let us now turn on the penalty term choice in (14). By the definition of $\sigma_{m'}$ in (12), it is smaller than the solution $\sigma'_{m'}$ of

$$\sigma = \sqrt{D_{m'}\left(1 + \frac{1}{\beta_{m'}}\log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right)} + \frac{\mathring{A}_{m'}}{\sigma}D_{m'}\left(1 + \frac{1}{\beta_{m'}}\log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right),$$

which satisfies

$$\sigma'_{m'} \geqslant \sqrt{\mathring{A}_{m'}D_{m'}(1 + \frac{1}{\beta_{m'}})}$$

and therefore

$$\begin{aligned} \sigma'_{m'} &\leqslant \sqrt{D_{m'}\left(1 + \frac{1}{\beta_{m'}}\log\left(\frac{\sqrt{2n\mathring{A}_{m'}}}{\sqrt{D_{m'}(1 + \frac{1}{\beta_{m'}})}} \vee e\right)\right)} \\ &\quad + \sqrt{\frac{\mathring{A}_{m'}D_{m'}}{1 + \frac{1}{\beta_{m'}}}}\left(1 + \frac{1}{\beta_{m'}}\underbrace{\log\left(\frac{\sqrt{2n\mathring{A}_{m'}}}{\sqrt{D_{m'}(1 + \frac{1}{\beta_{m'}})}} \vee e\right)}_{\mathbf{T}}\right) \\ &= \sqrt{D_{m'}(1 + \frac{1}{\beta_{m'}})}\left(\sqrt{\frac{1 + \frac{1}{\beta_{m'}}\mathbf{T}}{1 + \frac{1}{\beta_{m'}}}} + \sqrt{\mathring{A}_{m'}}\frac{1 + \frac{1}{\beta_{m'}}\mathbf{T}}{1 + \frac{1}{\beta_{m'}}}\right). \end{aligned}$$

Since $\mathbf{T} \geqslant 1$ by its definition and $\mathring{A}_{m'} \geqslant 1$ by assumption,

$$\begin{aligned} \sigma'_{m'} &\leqslant 2\sqrt{D_{m'}\mathring{A}_{m'}(1 + \frac{1}{\beta_{m'}})}\frac{1 + \frac{1}{\beta_{m'}}\mathbf{T}}{1 + \frac{1}{\beta_{m'}}} \\ &\leqslant 2\sqrt{\mathring{A}_{m'}D_{m'}}\left(\sqrt{1 + \frac{1}{\beta_{m'}}} \vee \frac{1 + \frac{1}{\beta_{m'}}\log\left(\sqrt{2n\mathring{A}_{m'}}\right)}{\sqrt{1 + \frac{1}{\beta_{m'}}}}\right). \end{aligned}$$

Because $\log(2n\mathring{A}_{m'}) \geqslant 1$,

$$\sigma'_{m'} \leqslant 2\sqrt{\frac{\mathring{A}_{m'}D_{m'}}{1 + \frac{1}{\beta_{m'}}}}\left(1 + \frac{1}{\beta_{m'}}\log\left(2n\mathring{A}_{m'}\right)\right)$$

$$\leqslant 2\sqrt{\mathring{A}_{m'}D_{m'}}\sqrt{1+\frac{1}{\beta_{m'}}\log\left(2n\mathring{A}_{m'}\right)}$$
$$\leqslant 4\sqrt{\mathring{A}_{m'}D_{m'}}\sqrt{1+\frac{1}{\beta_{m'}}\log\left(n\mathring{A}_{m'}\right)}.$$

There exist numerical constants C_{pen} and C such that if for all $m' \in \mathcal{M}'$,

$$\text{(17)} \qquad \mathrm{pen}(m') \geqslant \frac{C_{\mathrm{pen}}}{\kappa}\mathring{A}^2_{m'}F^{\infty}_{m'}\left(1+\frac{1}{\beta_{m'}}\right)\log(n\mathring{A}_{m'})^2\frac{D_{m'}}{n},$$

with the choice of $\eta_{m'}$ in (15) and (16), and taking into account that (17) ensures that the term $\mathrm{Res}(m')$ and $\mathrm{Res}(m)$ are absorbed by the penalty terms (recall with Lemma 11 that $\mathrm{Res}(m') = 6F^{\infty}_{m'}n^{-(1+\alpha)}$), (14) becomes, for all $x \geqslant 0$, with probability at least $1-q_n-24\log(n)\Sigma e^{-x}$, for all $m' \in \mathcal{M}'$,

$$\text{(18)} \qquad (1-\kappa)\mathbf{K}_n(p^{m'}_{\hat{\theta}^{m'}}) \leqslant (1+\kappa)\mathbf{K}_n(p^{m}_{\bar{\theta}^{m}}) + 2\mathrm{pen}(m) + \frac{C}{\kappa}\left[\mathring{A}_{m'}F^{\infty}_{m'} + \mathring{A}_{m}F^{\infty}_{m}\right]\frac{x}{n}.$$

Let us distinguish the final result between Theorem 2 and 4.

- For Theorem 2, (17) reads

$$\mathrm{pen}(m') \geqslant \frac{C_{\mathrm{pen}}}{\kappa}A^2_{m'}\log(\varepsilon^{-1})\left(1+\frac{1}{\beta_{m'}}\right)\log(nA_{m'})^2\frac{D_{m'}}{n}$$

  and (18) becomes: for all $x \geqslant 0$, with probability at least $1-24\log(n)\Sigma e^{-x}$, for all $m' \in \mathcal{M}'$,

$$(1-\kappa)\mathbf{K}_n(p^{m'}_{\hat{\theta}^{m'}}) \leqslant (1+\kappa)\mathbf{K}_n(p^{m}_{\bar{\theta}^{m}}) + 2\mathrm{pen}(m) + \frac{2C}{\kappa}(A_{m'}+A_m)\log(\varepsilon^{-1})\frac{x}{n}.$$

- For Theorem 4, up to a multiplicative constant, since $\mathring{A}_{m''} = A_{m''}\log n$, (17) reads

$$\mathrm{pen}(m') \geqslant \frac{(1+\alpha)C_{\mathrm{pen}}}{\kappa}A^2_{m'}B_{m'}(\log n)^3\left(1+\frac{1}{\beta_{m'}}\right)\log(nA_{m'})^2\frac{D_{m'}}{n}$$

  and (18) becomes: for all $x \geqslant 0$, with probability at least $1-2n^{-\alpha}-24\log(n)\Sigma e^{-x}$, for all $m' \in \mathcal{M}'$,

$$(1-\kappa)\mathbf{K}_n(p^{m'}_{\hat{\theta}^{m'}}) \leqslant (1+\kappa)\mathbf{K}_n(p^{m}_{\bar{\theta}^{m}}) + 2\mathrm{pen}(m) + \frac{(1+\alpha)C}{\kappa}(A_{m'}B_{m'}+A_mB_m)\frac{x(\log n)^2}{n}.$$

APPENDIX B: PROOF OF COROLLARIES 3 AND 5

Corollary 3 follows directly from the fact that $\mathbb{E}[Z] \leqslant a\int_{t\geqslant 0}\mathbb{P}(Z \geqslant at)dt$ for any positive random variable Z, and any $a>0$.

Under Assumptions 1bis, let us first show that $\mathbf{K}_n(p^{m'}_{\delta})$ is bounded for all δ and m' almost surely. By Assumption 1bis, almost surely, for any $m' \in \mathcal{M}$ and $\delta \in \Theta_{m'}$,

$$\mathbf{K}_n(p^{m'}_{\delta}) \leqslant \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left[\left|\log\frac{p^{\star}_t(X_t|\mathcal{F}_{t-1})}{p^{m'}_{\delta,t}(X_t|\mathcal{F}_{t-1})}\right|\,\middle|\,\mathcal{F}_{t-1}\right]$$
$$\leqslant \frac{1}{n}\sum_{t=1}^{n}\left(B_{m'} + B_{m'}\int_1^{+\infty}\mathbb{P}\left(\left|\log\frac{p^{\star}_t(X_t|\mathcal{F}_{t-1})}{p^{m'}_{\delta,t}(X_t|\mathcal{F}_{t-1})}\right| \geqslant B_{m'}y\,\middle|\,\mathcal{F}_{t-1}\right)dy\right)$$
$$\text{(19)} \qquad \leqslant 2B_{m'}.$$

To conclude, assume that there exist $A(n), B(n)$ such that $\sup_{m\in\mathcal{M}} B_m \leqslant B(n)$, $\sup_{m\in\mathcal{M}} A_m \leqslant A(n)$, so that by Theorem 4, with probability at least $1 - 2n^{-\alpha} - 24\log(n)\Sigma e^{-x}$,

$$(1-\kappa)\mathbf{K}_n(\tilde{p}) \leqslant \inf_{m\in\mathcal{M}} \left((1+\kappa) \inf_{\theta\in\Theta^{D_m}} \mathbf{K}_n(p_\theta^m) + 2\mathrm{pen}(m) \right) + \frac{2(1+\alpha)C'}{\kappa} A(n)B(n)^2 \frac{(\log n)^2 x}{n},$$

and use that for any random variable Z such that $Z \leqslant M$ a.s. for some constant $M > 0$, $\mathbb{E}[Z] \leqslant \int_{t=0}^{M} \mathbb{P}(Z \geqslant t)dt$, so that for all $\kappa \in (0,1]$,

$$(1-\kappa)\mathbb{E}\left[\mathbf{K}_n(p_{\hat{\theta}^{\hat{m}}}^{\hat{m}})\right] \leqslant \mathbb{E}\left[\inf_{m\in\mathcal{M}} \left((1+\kappa) \inf_{\theta\in\Theta^{D_m}} \mathbf{K}_n(p_\theta^m) + 2\mathrm{pen}(m) \right)\right] + 4B(n)n^{-\alpha} + \frac{48(1+\alpha)C'}{\kappa} \Sigma A(n)B(n)^2 \frac{(\log n)^3}{n},$$

and the last term dominates the second to last one when $\alpha \geqslant 1$.

APPENDIX C: PROOF OF THE LEMMAS

C.1. Proof of Lemma 11. For any $m' \in \mathcal{M}$ and $\theta \in \Theta_{m'}$

$$|\mathbf{K}_n(p_\theta^{m'}) - \widetilde{\mathbf{K}}_n(p_\theta^{m'})| \leqslant \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[\left| \log \frac{p_t^\star(X_t)}{p_{\theta,t}^{m'}(X_t)} \right| \mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1} \right].$$

Let $y \geqslant 0$ to be determined later. Let $Y_t = \left|\log \frac{p_t^\star(X_t)}{p_{\theta,t}^{m'}(X_t)}\right|$. Then,

$$\begin{aligned}
\mathbb{E}\left[Y_t \mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1}\right] &= \mathbb{E}\left[Y_t \mathbf{1}_{f_A^t(X_t)\geqslant 1} \mathbf{1}_{B_{m'}^{-1}Y_t > y} \middle| \mathcal{F}_{t-1}\right] + \mathbb{E}\left[Y_t \mathbf{1}_{f_A^t(X_t)\geqslant 1} \mathbf{1}_{B_{m'}^{-1}Y_t \leqslant y} \middle| \mathcal{F}_{t-1}\right] \\
&\leqslant \mathbb{E}\left[Y_t \mathbf{1}_{B_{m'}^{-1}Y_t > y} \middle| \mathcal{F}_{t-1}\right] + B_{m'} y \mathbb{E}\left[\mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1}\right] \\
&= B_{m'}\left(\mathbb{E}\left[B_{m'}^{-1} Y_t \mathbf{1}_{B_{m'}^{-1}Y_t > y} \middle| \mathcal{F}_{t-1}\right] + y\mathbb{E}\left[\mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1}\right] \right).
\end{aligned}$$

By Assumption 1bis, conditionally to $\mathcal{F}_{t-1}$, $B_{m'}^{-1} Y_t$ is stochastically dominated by $1 \vee E$ where E is an exponentially distributed random variable with parameter 1. Therefore, for $y \geqslant 1$, a.s.,

$$\mathbb{E}\left[B_{m'}^{-1} Y_t \mathbf{1}_{B_{m'}^{-1}Y_t > y} \middle| \mathcal{F}_{t-1}\right] \leqslant \mathbb{E}[(E \vee 1)\mathbf{1}_{E\vee 1 \geqslant y} | \mathcal{F}_{t-1}] = \mathbb{E}[E \mathbf{1}_{E \geqslant y} | \mathcal{F}_{t-1}] \leqslant (1+y)e^{-y}.$$

Finally, since $\mathbb{P}(f_A^t(X_t) \geqslant 1 \,|\, \mathcal{F}_{t-1}) \leqslant 2n^{-(1+\alpha)}$ a.s. by assumption, it holds almost surely that

$$\begin{aligned}
\mathbb{E}\left[Y_t \mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1}\right] &\leqslant B_{m'}\left((1+y)e^{-y} + y\mathbb{P}\left(f_A^t(X_t) \geqslant 1 \middle| \mathcal{F}_{t-1}\right) \right) \\
&\leqslant B_{m'}\left((1+y)e^{-y} + \frac{2y}{n^{1+\alpha}} \right).
\end{aligned}$$

Pick $y = \log\left(\frac{n^{1+\alpha}}{2}\right) \geqslant 1$. Thus,

$$\begin{aligned}
\mathbb{E}\left[Y_t \mathbf{1}_{f_A^t(X_t)\geqslant 1} \middle| \mathcal{F}_{t-1}\right] &\leqslant \frac{2B_{m'}}{n^{1+\alpha}}\left(1 + 2\log\left(\frac{n^{1+\alpha}}{2}\right)\right) \\
&\leqslant 6B_{m'}(1+\alpha)\frac{\log n}{n^{1+\alpha}}.
\end{aligned}$$

Therefore, $|\mathbf{K}_n(p_\theta^{m'}) - \widetilde{\mathbf{K}}_n(p_\theta^{m'})| \leqslant 6B_{m'}(1+\alpha)\frac{\log n}{n^{1+\alpha}}$.

C.2. Proof of Lemma 12. Fix $m' \in \mathcal{M}$. Recall that for any $\delta, \eta \in \Theta_{m'}$,

$$\nu(g_\delta^{m'}) - \nu(g_\eta^{m'}) = \frac{1}{n}\sum_{t=1}^n \int \log\left(\frac{p_{\eta,t}^{m'}(x_t|\mathcal{F}_{t-1})}{p_{\delta,t}^{m'}(x_t|\mathcal{F}_{t-1})}\right)\mathbf{1}_{f_A^t(x_t)<1}(d\delta_{X_t}(x_t) - p_t^\star(x_t|\mathcal{F}_{t-1})d\mu(x_t)),$$

where δ_a is the Dirac measure in a.

We extend $\Theta_{m'}$ into $\Theta_{m'} \cup \{\star\}$ by defining $p_\star^{m'} = p^\star$, so that when $\eta = \star$, $\nu(g_\eta^{m'}) = 0$ and the formula above becomes $\nu(g_\delta^{m'})$. We want to control this uniformly over $\delta, \eta \in \Theta_{m'} \cup \{\star\}$.

Fix $\delta, \eta \in \Theta_{m'} \cup \{\star\}$. For any $t \in [n]$, let

$$\Delta_t = \int \log\left(\frac{p_{\eta,t}^{m'}(x_t|\mathcal{F}_{t-1})}{p_{\delta,t}^{m'}(x_t|\mathcal{F}_{t-1})}\right)\mathbf{1}_{f_A^t(x_t)<1}(d\delta_{X_t}(x_t) - p_t^\star(x_t|\mathcal{F}_{t-1})d\mu(x_t)).$$

For any $t \in [n+1]$, let $M_t = \sum_{s=1}^{t-1}\Delta_s$ (in particular, $M_1 = 0$), so that $\frac{1}{n}M_{n+1} = \nu(g_\delta^{m'}) - \nu(g_\eta^{m'})$.

$(M_t)_{t\geqslant 1}$ is a $(\sigma(\mathcal{F}_{t-1}))_{t\geqslant 1}$-martingale. For $\ell \geqslant 2$, let $C_1^\ell = 0$ and for $t \geqslant 2$, let

$$C_t^\ell = \sum_{s=1}^{t-1}\mathbb{E}[\Delta_s^\ell \,|\, \mathcal{F}_{t-1}].$$

Note that for all $s \in [n]$,

$$|\Delta_s| \leqslant 2\int \left|\log\frac{p_{\eta,s}^{m'}(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})}\right|\mathbf{1}_{f_A^s(x_s)<1}\frac{d\delta_{X_s}(x_s) + p_s^\star(x_s|\mathcal{F}_{s-1})d\mu(x_s)}{2},$$

so that by convexity of $x \mapsto |x|^\ell$,

$$|C_t^\ell| \leqslant \sum_{s=1}^{t-1}\mathbb{E}\left[2^\ell\int\left|\log\frac{p_{\eta,s}^{m'}(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})}\right|^\ell \mathbf{1}_{f_A^s(x_s)<1}\frac{d\delta_{X_s}(x_s) + p_s^\star(x_s|\mathcal{F}_{s-1})d\mu(x_s)}{2}\,\middle|\,\mathcal{F}_{s-1}\right]$$

$$= 2^\ell\sum_{s=1}^{t-1}\int\left|\log\frac{p_{\eta,s}^{m'}(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})}\right|^\ell \mathbf{1}_{f_A^s(x_s)<1}p_s^\star(x_s|\mathcal{F}_{s-1})d\mu(x_s). \tag{20}$$

Let

$$R_{\infty,n}(\delta,\eta) = \max_{1\leqslant s\leqslant n}\sup_{x\in\mathcal{X}}\left(\left|\log\frac{p_{\eta,s}^{m'}(x|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x|\mathcal{F}_{s-1})}\right|\mathbf{1}_{f_A^s(x)<1}\right) \tag{21}$$

and

$$R_{2,n}(\delta,\eta)^2 = 2\sum_{s=1}^n\mathbb{E}\left[\left(\log\frac{p_{\eta,s}^{m'}(X_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(X_s|\mathcal{F}_{s-1})}\right)^2\mathbf{1}_{f_A^s(X_s)<1}\,\middle|\,\mathcal{F}_{s-1}\right].$$

Under Assumptions 1, 2 and 4, almost surely, for all $x \in \mathcal{X}$ and $s \in [n]$, $\mathbf{1}_{f_A^s(x)<1} = 1$ and therefore

$$R_{\infty,n}(\delta,\eta) = \max_{1\leqslant s\leqslant n}\sup_{x\in\mathcal{Q}}\left|\log\frac{p_{\theta,s}^m(x)}{p_{\delta,s}^m(x)}\right| \quad \text{almost surely.}$$

In what follows, when using the notation $R_\infty(\delta,\eta)$, we will actually mean the right hand term, which is $\mathcal{F}_{n-1}$-measurable as a supremum of a countable family of measurable functions.

Secondly, under Assumptions 1bis, 2bis and 4bis, for all $x \in \mathcal{X}$ and $s \in [n]$, with O_x the open set of Assumption 4bis, the restriction of the function f_A^s to $O_x \cup \{x\}$ is continuous at x. Therefore, for all $s \in [n]$, if $f_A^s(x) < 1$ for all $s \in [n]$, the function

$$y \in O_x \cup \{x\} \mapsto \left| \log \frac{p_{\eta,s}^{m'}(y|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(y|\mathcal{F}_{s-1})} \right| \mathbf{1}_{f_A^s(y)<1}$$

is continuous at x. Thus, for any dense countable subset $\mathcal{Q}$ of $\mathcal{X}$ and any $y \in \mathcal{X}$ and $s \in [n]$, if $f_A^s(y) < 1$,

$$\sup_{x \in \mathcal{Q}} \left(\left| \log \frac{p_{\eta,s}^{m'}(x|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x|\mathcal{F}_{s-1})} \right| \mathbf{1}_{f_A^s(x)<1} \right) \geqslant \left| \log \frac{p_{\eta,s}^{m'}(y|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(y|\mathcal{F}_{s-1})} \right| \mathbf{1}_{f_A^s(y)<1},$$

and if $f_A^s(y) \geqslant 1$, then the right hand side is zero and so the inequality still holds. Therefore, (21) can be rewritten as

$$R_{\infty,n}(\delta,\eta) = \max_{1 \leqslant s \leqslant n} \sup_{x \in \mathcal{Q}} \left(\left| \log \frac{p_{\eta,s}^{m'}(x|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x|\mathcal{F}_{s-1})} \right| \mathbf{1}_{f_A^s(x)<1} \right),$$

for any dense countable subset $\mathcal{Q}$ of $\mathcal{X}$. Since this is a supremum of a countable family of $\mathcal{F}_{n-1}$-measurable functions, $R_{\infty,n}(\delta,\eta)$ is $\mathcal{F}_{n-1}$-measurable.

Injecting these quantities into (20) shows that

$$|C_t^\ell| \leqslant 2^{\ell-1} R_{\infty,n}(\delta,\eta)^{\ell-2} R_{2,n}(\delta,\eta)^2,$$

and since $2^{\ell-1} \leqslant \ell!$ for all $\ell \geqslant 2$,

$$|C_{n+1}^\ell| \leqslant \frac{\ell!}{2} R_{2,n}(\delta,\eta)^2 R_{\infty,n}(\delta,\eta)^{\ell-2}. \tag{22}$$

The next proposition is a direct consequence of Lemma 3.3 in [36] and can be found in [4],

PROPOSITION 14. *Let $(M_t)_{t\geqslant 0}$ be a $(\mathcal{F}_t)_{t\geqslant 0}$-martingale with $M_0 = 0$. Let $t \geqslant 1$ and assume that there exists nonnegative random variables R_2 and R_∞ such that for all $\ell \geqslant 2$,*

$$\sum_{s=1}^{t} \mathbb{E}\left[(M_s - M_{s-1})^\ell \,\middle|\, \mathcal{F}_{s-1} \right] \leqslant \frac{\ell!}{2} R_2^2 R_\infty^{\ell-2}.$$

Then, for all $\sigma, \sigma', x \geqslant 0$,

$$\mathbb{P}\left(M_t \geqslant \sigma\sqrt{2x} + \sigma' x,\ R_2 \leqslant \sigma \text{ and } R_\infty \leqslant \sigma' \right) \leqslant e^{-x}. \tag{23}$$

PROOF OF PROPOSITION 14. Let $C_0^\ell = 0$ and $C_s^\ell = \sum_{u=1}^s \mathbb{E}[(M_u - M_{u-1})^\ell \,|\, \mathcal{F}_{u-1}]$ for all $t \geqslant s \geqslant 1$ and $\ell \geqslant 2$. Lemma 3.3 of [36] gives that for all $\lambda > 0$, the sequence $(\mathcal{E}_s)_{s\geqslant 0}$ defined by

$$\mathcal{E}_s = \exp\left(\lambda M_s - \sum_{\ell \geqslant 2} \frac{\lambda^\ell}{\ell!} C_s^\ell \right)$$

is a supermartingale. In particular, $\mathbb{E}(\mathcal{E}_t) \leqslant \mathbb{E}(\mathcal{E}_0) = 1$. Therefore,

$$\forall \lambda \geqslant 0 \qquad \mathbb{E}\left[\exp\left(\lambda M_t - \frac{(\lambda R_2)^2}{2} \sum_{k\geqslant 0} (\lambda R_\infty)^k \right) \right] \leqslant 1. \tag{24}$$

By (24), for all $\sigma, \sigma', \lambda \geqslant 0$,

$$\mathbb{E}\left[\exp\left(\lambda M_t - \frac{(\lambda R_2)^2}{2}\sum_{k\geqslant 0}(\lambda R_\infty)^k\right)\mathbf{1}_{R_2\leqslant\sigma}\mathbf{1}_{R_\infty\leqslant\sigma'}\right] \leqslant 1,$$

and therefore for all $\lambda \in [0, (\sigma')^{-1})$,

$$\mathbb{E}[\exp(\lambda M_t)\mathbf{1}_{R_2\leqslant\sigma}\mathbf{1}_{R_\infty\leqslant\sigma'}] \leqslant \exp\left(\frac{\lambda^2\sigma^2}{2(1-\lambda\sigma')}\right). \tag{25}$$

Finally, for $y \geqslant 0$, for all $\lambda \geqslant 0$,

$$\begin{aligned}\mathbb{P}\left(M_t \geqslant y,\ R_2 \leqslant \sigma \text{ and } R_\infty \leqslant \sigma'\right) &= \mathbb{P}\left(e^{\lambda M_t} \geqslant e^{\lambda y},\ R_2 \leqslant \sigma \text{ and } R_\infty \leqslant \sigma'\right) \\ &= \mathbb{P}\left(e^{\lambda M_t}\mathbf{1}_{R_2\leqslant\sigma}\mathbf{1}_{R_\infty\leqslant\sigma'} \geqslant e^{\lambda y}\right)\end{aligned}$$

where the last equality holds because $e^{\lambda M_t}$ and $e^{\lambda y}$ are positive. Therefore, by the Markov inequality and (25), for all $\lambda \in [0, (\sigma')^{-1})$,

$$\mathbb{P}\left(M_t \geqslant y,\ R_2 \leqslant \sigma \text{ and } R_\infty \leqslant \sigma'\right) \leqslant \exp\left(\frac{\lambda^2\sigma^2}{2(1-\lambda\sigma')} - \lambda y\right).$$

Classic Bernstein inequality's proofs show that an optimal choice of λ leads to

$$\mathbb{P}\left(M_t \geqslant y,\ R_2 \leqslant \sigma \text{ and } R_\infty \leqslant \sigma'\right) \leqslant \exp\left(-\frac{\sigma^2}{(\sigma')^2}h\left(\frac{\sigma' y}{\sigma^2}\right)\right),$$

where h is the function defined by $h(u) = 1 + u - \sqrt{1+2u}$ for all $u \geqslant 0$. Since, its inverse function h^{-1} is such that $h^{-1}(u) = u + \sqrt{2u}$ for all $u \geqslant 0$, choosing y so that

$$\frac{\sigma^2}{(\sigma')^2}h\left(\frac{\sigma' y}{\sigma^2}\right) = x$$

leads to

$$y = \frac{\sigma^2}{\sigma'}h^{-1}\left(\frac{(\sigma')^2 x}{\sigma^2}\right) = \sigma' x + \sigma\sqrt{2x}.$$

Hence the result. □

Applying Proposition 14 to the martingale M_t ensures that for all $\sigma, \sigma', x \geqslant 0$ and for all $\delta, \eta \in \Theta_{m'} \cup \{\star\}$,

$$\mathbb{P}\left(n\nu(g_\delta^{m'}) - n\nu(g_\eta^{m'}) \geqslant \sigma\sqrt{2x} + \sigma' x, R_2(\delta,\eta) \leqslant \sigma \text{ and } R_\infty(\delta,\eta) \leqslant \sigma'\right) \leqslant e^{-x}.$$

From this control of the increments of the process ν, we aim to apply the following slight modification of Theorem 5 of [4], to account for a Hölder relation between R_2, R_∞ and the norms over $\Theta_{m'}$.

THEOREM 15. *Let $(Y_t)_{t\in S}$ be a family of real-valued random variables, indexed by a countable subset S of a linear space E of finite dimension D. Let R_2 and R_∞ be two non-negative random functions on $S \times S$ that satisfy the triangle inequality. Assume that the increments $Y_s - Y_u$ satisfy*

(26) $\forall \sigma, \sigma', x \geqslant 0, \quad \forall s, u \in S,$

$$\mathbb{P}(Y_s - Y_u \geqslant \sigma\sqrt{2x} + \sigma' x, R_2(s,u) \leqslant \sigma \textit{ and } R_\infty(s,u) \leqslant \sigma') \leqslant e^{-x}.$$

Assume that there exists two deterministic seminorms N_2 and N_∞ on E, a positive constant β and nonnegative constants v, w, such that for all $s, u \in S$,

$$\begin{cases} R_2(s,u) \leqslant N_2(s-u)^\beta \leqslant v^\beta, \\ R_\infty(s,u) \leqslant N_\infty(s-u)^\beta \leqslant w^\beta. \end{cases} \tag{27}$$

Let $c \geqslant 0$. Define for all $x, \sigma \geqslant 0$,

$$\Psi(\sigma, x) = 30\Bigg[(\sigma \wedge v^\beta)\sqrt{x + D + \frac{D}{\beta}\log\left(\frac{v^\beta \vee (cw^\beta)}{\sigma} \vee e\right)} + (\frac{\sigma}{c} \wedge w^\beta)\left(x + D + \frac{D}{\beta}\log\left(\frac{v^\beta \vee (cw^\beta)}{\sigma} \vee e\right)\right)\Bigg],$$

with the convention $\frac{\sigma}{0} = +\infty$. Finally, define for all $t, t_0 \in S$, $\Delta_c(t,t_0) = R_2(t,t_0) \vee (cR_\infty(t,t_0))$. Then, for any $t_0 \in S$, $x \geqslant 0$ and $\sigma > 0$,

$$\mathbb{P}\left(\sup_{t\in S} \frac{Y_t - Y_{t_0}}{\Delta_c(t,t_0)^2 + \sigma^2} \geqslant 4\sigma^{-2}\Psi(\sigma, x)\right) \leqslant 2\left(\log\left(\frac{v^\beta \vee (cw^\beta)}{\sigma}\right) \vee 0 + 1\right) e^{-x}. \tag{28}$$

The salient differences in the proof of this Theorem compared to Theorem 5 of [4] are detailed in Section C.4.

We take S as either $\Theta_{m'}$ or a dense countable subset of $\Theta_{m'}$; the two choices are equivalent under Assumption 3. Such a dense countable subset always exists since $\Theta_{m'}$ is parametric. Let us now check the assumptions of the theorem.

By definition,

$$\begin{cases} \forall \delta, \eta \in \Theta_{m'} \cup \{\star\}, \quad R_{2,n}(\delta,\eta) \leqslant F^{\infty}_{m'}\sqrt{2n}, \\ \qquad\qquad \forall \delta, \eta \in \Theta_{m'}, \quad R_{2,n}(\delta,\eta) \leqslant F^{\text{höl}}_{m'}\|\delta - \eta\|_{m'}^{\beta_{m'}}\sqrt{2n}. \end{cases} \tag{29}$$

Since $R_{2,n}(\delta,\eta)$ is the Euclidean norm of the vector whose coordinate $s \in [n]$ is the $\mathbf{L}^2(\mathcal{X}, p^\star_s(\cdot|\mathcal{F}_{s-1})d\mu)$ distance between $\log p^{m'}_{\eta,s}(\cdot|\mathcal{F}_{s-1})\mathbf{1}_{f^s_A(\cdot)<1}$ and $\log p^{m'}_{\delta,s}(\cdot|\mathcal{F}_{s-1})\mathbf{1}_{f^s_A(\cdot)<1}$, it satisfies the triangular inequality: for all $\eta, \delta, \theta \in \Theta_{m'} \cup \{\star\}$,

$$R_{2,n}(\delta,\eta) \leqslant R_{2,n}(\delta,\theta) + R_{2,n}(\theta,\eta). \tag{30}$$

Likewise,

$$\begin{cases} \forall \delta, \eta \in \Theta_{m'} \cup \{\star\}, \quad R_{\infty,n}(\delta,\eta) \leqslant F^{\infty}_{m'}, \\ \qquad\qquad \forall \delta, \eta \in \Theta_{m'}, \quad R_{\infty,n}(\delta,\eta) \leqslant F^{\text{höl}}_{m'}\|\delta - \eta\|_{m'}^{\beta_{m'}}, \end{cases} \tag{31}$$

and for all $\eta, \delta, \theta \in \Theta_{m'} \cup \{\star\}$

$$R_{\infty,n}(\delta,\eta) \leqslant R_{\infty,n}(\delta,\theta) + R_{\infty,n}(\theta,\eta). \tag{32}$$

Identify $\Theta_{m'} \cup \{\star\}$ with the subset $\widetilde{\Theta}_{m'}$ of the vector space $\mathbb{R}^{D_{m'}+1}$ of generic element (δ, u) with $\delta \in \mathbb{R}^{D_{m'}}$ and $u \in \mathbb{R}$, defined as

$$\widetilde{\Theta}_{m'} = \{(\theta, 0) : \theta \in \Theta_{m'}\} \cup \{(\bar{\theta}^{m'}, 1)\},$$

for some fixed $\bar{\theta}^{m'} \in \Theta_{m'}$. Endow the vector space $\mathbb{R}^{D_{m'}+1}$ with the norms

$$\widetilde{N}_\infty((\delta,u)) = (F^{\text{höl}}_{m'}\|\delta\|_{m'}^{\beta_{m'}} + F^{\infty}_{m'}|u|^{\beta_{m'}})^{1/\beta_{m'}} \quad \text{and} \quad \widetilde{N}_2((\delta,u)) = (2n)^{1/(2\beta_{m'})}\widetilde{N}_\infty((\delta,u)).$$

Under Assumption 2bis, for any $\delta, \eta \in \Theta_{m'}$, by (29) and (31),

$$\begin{cases} R_{2,n}(\delta,\eta) \leqslant F^{\text{höl}}_{m'} \|\delta-\eta\|_{m'}^{\beta_{m'}} \sqrt{2n} = \widetilde{N}_2((\delta,0)-(\eta,0))^{\beta_{m'}} \\ \quad R_{\infty,n}(\delta,\eta) \leqslant F^{\text{höl}}_{m'} \|\delta-\eta\|_{m'}^{\beta_{m'}} = \widetilde{N}_\infty((\delta,0)-(\eta,0))^{\beta_{m'}}, \end{cases}$$

and both inequalities extend to $\eta = \star$ since, for $R_{2,n}$,

$$\begin{aligned} R_{2,n}(\delta,\star) &\leqslant F^{\infty}_{m'}\sqrt{2n} \\ &\leqslant \left[(F^{\text{höl}}_{m'})^{1/\beta_{m'}} \|\delta - \bar\theta^{m'}\|_{m'} (2n)^{1/(2\beta_{m'})} + (F^{\infty}_{m'})^{1/\beta_{m'}} (2n)^{1/(2\beta_{m'})} \right]^{\beta_{m'}} \\ &= \widetilde{N}_2((\delta,0)-(\bar\theta^{m'},1))^{\beta_{m'}} \end{aligned}$$

and likewise for $\delta = \star$ and for $R_{\infty,n}$. Recall that we defined

$$\mathring{A}_{m'} = F^{\text{höl}}_{m'} M^{\beta_{m'}}_{m'} + F^{\infty}_{m'} \quad \text{and} \quad v_{m'} = \mathring{A}_{m'}\sqrt{2n},$$

in (11), so that $\widetilde{N}_2(\delta-\eta)^{\beta_{m'}} \leqslant v_{m'}$ and $\widetilde{N}_\infty(\delta-\eta)^{\beta_{m'}} \leqslant \mathring{A}_{m'}$ for all $\delta,\eta \in \widetilde\Theta_{m'}$.

We may now apply Theorem 15 to the process $Y_\delta = n\nu(g^{m'}_\delta)$ indexed by $\delta \in \Theta_{m'} \cup \{\star\}$, with $c = 0$: for all $\sigma > 0$ and $x \geqslant 0$, let

$$\Psi(\sigma,x) = 30\left[(\sigma \wedge v_{m'})\sqrt{x + D_{m'} + \frac{D_{m'}}{\beta_{m'}} \log\left(\frac{v_{m'}}{\sigma} \vee e\right)} + \mathring{A}_{m'}\left(x + D_{m'} + \frac{D_{m'}}{\beta_{m'}} \log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right)\right].$$

Then, for all $\theta \in \Theta_{m'} \cup \{\star\}$, $\sigma > 0$ and $x \geqslant 0$,

$$\mathbb{P}\left(\sup_{\delta \in \Theta_{m'} \cup \{\star\}} \frac{Y_\delta - Y_\theta}{R_{2,n}(\delta,\theta)^2 + \sigma^2} \geqslant 4\sigma^{-2}\Psi(\sigma, x + D_{m'}) \right) \leqslant \left(2\log\left(\frac{v_{m'}}{\sigma}\right) \vee 0 + 1\right) e^{-(x+D_{m'})}.$$

In particular, by taking the union bound over $m' \in \mathcal{M}$ for $\theta = \star$, for any family of positive numbers $(y_{m'})_{m' \in \mathcal{M}}$, with probability at least $1 - e^{-x} \sum_{m' \in \mathcal{M}} \left(2\log\left(\frac{v_{m'}}{y_{m'}}\right) \vee 0 + 1\right) e^{-D_{m'}}$, for all $m' \in \mathcal{M}$,

$$\begin{aligned} \sup_{\delta \in \Theta_{m'}} \left(\frac{n\nu(g^{m'}_\delta)}{R_{2,n}(\delta,\star)^2 + y^2_{m'}} \right) &\leqslant \frac{120}{y^2_{m'}} \left[y_{m'}\sqrt{x + D_{m'}} + (y_{m'} \wedge v_{m'}) \sqrt{D_{m'}\left(1 + \frac{1}{\beta_{m'}} \log\left(\frac{v_{m'}}{y_{m'}} \vee e\right)\right)} \right. \\ &\qquad \left. + \mathring{A}_{m'}(x + D_{m'}) + \mathring{A}_{m'} D_{m'} \left(1 + \frac{1}{\beta_{m'}} \log\left(\frac{v_{m'}}{y_{m'}} \vee e\right)\right)\right]. \end{aligned}$$

For each $m' \in \mathcal{M}$, let $\sigma_{m'}$ be the solution of the equation

$$\sigma = \left(1 \wedge \frac{v_{m'}}{\sigma}\right) \sqrt{D_{m'}\left(1 + \frac{1}{\beta_{m'}} \log\left(\frac{v_{m'}}{\sigma} \vee e\right)\right)}$$

$$+\frac{\mathring{A}_{m'}}{\sigma}D_{m'}\left(1+\frac{1}{\beta_{m'}}\log\left(\frac{v_{m'}}{\sigma}\vee e\right)\right)$$

which exists since the right hand side is positive and non-increasing on $(0,+\infty)$. For any family $(y_{m'})_{m'\in\mathcal{M}}$ such that $y_{m'}\geqslant\sigma_{m'}$ for all $m'\in\mathcal{M}$, for any $x\geqslant 0$, it holds with probability at least $1-e^{-x}\sum_{m'\in\mathcal{M}}\left(2\log\left(\frac{v_{m'}}{y_{m'}}\right)\vee 0+1\right)e^{-D_{m'}}$ that for all $m'\in\mathcal{M}$,

$$\sup_{\delta\in\Theta_{m'}}\left(\frac{n\nu(g_\delta^{m'})}{R_{2,n}(\delta,\star)^2+y_{m'}^2}\right)\leqslant\frac{120}{y_{m'}}\left(\sigma_{m'}+\sqrt{x+D_{m'}}+\frac{\mathring{A}_{m'}}{y_{m'}}(x+D_{m'})\right).$$

Let $\eta\in(0,1)$ and $x\geqslant 0$ and fix for each $m'\in\mathcal{M}$

$$y_{m'}=\eta^{-1}\sqrt{\sigma_{m'}^2+x+D_{m'}}\ ,$$

then with probability at least $1-e^{-x}\sum_{m'\in\mathcal{M}}\left(2\log\left(\frac{v_{m'}}{y_{m'}}\right)\vee 0+1\right)e^{-D_{m'}}$, for all $m'\in\mathcal{M}$,

$$\sup_{\delta\in\Theta_{m'}}\left(\frac{\nu(g_\delta^{m'})}{\frac{1}{n}R_{2,n}(\delta,\star)^2+\frac{1}{n}y_{m'}^2}\right)\leqslant 120(2\eta+\eta^2\mathring{A}_{m'}),$$

where

- the first term on the right hand side is due to the concavity of $x\in(0,+\infty)\mapsto\sqrt{x}$,
- the second term on the right hand side holds because $x+D_{m'}\leqslant x+D_{m'}+\sigma_{m'}^2=\eta^2y_{m'}^2$.

By definition $y_{m'}\geqslant\eta^{-1}\sqrt{D_{m'}}\geqslant 1$, and $v_{m'}=\mathring{A}_{m'}\sqrt{2n}\geqslant e$. Therefore,

$$2\log\left(\frac{v_{m'}}{y_{m'}}\right)\vee 0+1\leqslant 3\log v_{m'}.$$

The control of $-\nu(g_\delta^{m'})$ is identical, hence we may control all $|\nu(g_\delta^{m'})|$ with probability at least $1-6e^{-x}\sum_{m'\in\mathcal{M}}\log(v_{m'})e^{-D_{m'}}$ by union bound. To conclude the proof of the Lemma, note that $\frac{1}{n}R_{2,n}(\delta,\star)^2=2\mathbf{V}_n(p_\delta)$.

C.3. Proof of Lemma 13. Let's begin with a result proved in [46] which we slightly adapt to our situation. It is proved at the end of this section.

LEMMA 16 (Adaptation of Lemma 4 in [46]). *For any probability measures $\mathbb{P}$ and $\mathbb{Q}$ with densities p and q, and any $\lambda\in(0,1/2]$,*

$$\mathbb{P}\left\{\left(\log\frac{p}{q}\right)^2\mathbf{1}_{\left|\log\left(\frac{p}{q}\right)\right|\leqslant\log\left(\frac{1}{\lambda}\right)}\right\}\leqslant 8\left(1+\left(\log\frac{1}{\lambda}\right)^2\right)\mathbb{P}\left\{\left(\frac{q^{1/2}}{p^{1/2}}-1\right)^2\mathbf{1}_{\left|\log\left(\frac{p}{q}\right)\right|\leqslant\log\left(\frac{1}{\lambda}\right)}\right\}.$$

Let us apply Lemma 16 to $p=p_s^\star$ and $q=p_{\delta,s}^{m'}$ with $\lambda=\exp(-F_{m'}^\infty)$ for all $s\in[n]$ and $\delta\in\Theta_{m'}$. By definition of $F_{m'}^\infty$, $\lambda\leqslant 1/2$, so that for all $s\in[n]$ and $\delta\in\Theta_{m'}$,

$$\int\left|\log\frac{p_s^\star(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})}\right|^2\mathbf{1}_{f_A^s(X_s)<1}\,p_s^\star(x_s|\mathcal{F}_{s-1})d\mu(x_s)$$

$$\leqslant \int \left| \log \frac{p_s^\star(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})} \right|^2 \mathbf{1}_{\left|\log \frac{p_s^\star(x_s|\mathcal{F}_{s-1})}{p_{\delta,s}^{m'}(x_s|\mathcal{F}_{s-1})}\right| < F_{m'}^\infty} p_s^\star(x_s|\mathcal{F}_{s-1}) d\mu(x_s)$$

$$\leqslant 16 h^2 \left(p_s^\star(\cdot|\mathcal{F}_{s-1}), p_{\delta,s}^{m'}(\cdot|\mathcal{F}_{s-1}) \,\middle|\, \mathcal{F}_{s-1} \right) (1 + (F_{m'}^\infty)^2),$$

where

$$h^2 \left(p_t(\cdot|\mathcal{F}_{t-1}), q_t(\cdot|\mathcal{F}_{t-1}) \,\middle|\, \mathcal{F}_{t-1} \right) = \frac{1}{2} \int \left(\sqrt{p_t(x_t|\mathcal{F}_{t-1})} - \sqrt{q_t(x_t|\mathcal{F}_{t-1})} \right)^2 d\mu(x_t).$$

Let us recall a classical relation between the Hellinger distance and the Kullback-Leibler divergence, see for instance in [42, Lemma 7.23]: for any probability measures P and Q,

$$2h^2(P,Q) \leqslant \mathrm{KL}(P,Q).$$

Applying this inequality to the probability measures $P = p_t^\star(\cdot|X_1^{t-1})$ and $Q = p_{t,\hat{\theta}^{m'}}^{m'}(\cdot|X_1^{t-1})$ conditionally to X_1^{t-1} for all $t \in [n]$ shows that

$$\mathbf{V}_n(p_\delta^{m'}) \leqslant 8(1 + (F_{m'}^\infty)^2)\mathbf{K}_n(p_\delta^{m'}) \leqslant 16 (F_{m'}^\infty)^2 \mathbf{K}_n(p_\delta^{m'})$$

since $F_{m'}^\infty \geqslant 1$.

PROOF OF LEMMA 16. The proof follows exactly the same steps as [46]. Let $r : (0,+\infty) \to \mathbb{R}$ be the function defined implicitly by

$$\log(x) = 2(x^{1/2} - 1) - r(x)(x^{1/2} - 1)^2.$$

The function r is non-negative, decreasing, and $r(x) \leqslant 2\log(1/x)$ for all $x \in (0, 1/2]$ (see e.g. [34]). Let $\lambda \in (0, 1/2]$. Since for any $x \geqslant 1$, $|\log(x)| \leqslant 2|x^{1/2} - 1|$,

$$\mathbb{P}\left\{ \left(\log \frac{p}{q} \right)^2 \mathbf{1}_{1 \leqslant \frac{q}{p} \leqslant \frac{1}{\lambda}} \right\} \leqslant 4\mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{1 \leqslant \frac{q}{p} \leqslant \frac{1}{\lambda}} \right\}.$$

Moreover, by definition of r,

$$\begin{aligned}
&\mathbb{P}\left\{ \left(\log \frac{p}{q} \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} \\
&\leqslant 8\mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} + 2\mathbb{P}\left\{ r^2\left(\frac{q}{p}\right) \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^4 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} \\
&\leqslant 8\mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} + 2r^2(\lambda)\mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} \\
&\leqslant 8\mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\} + 8\left(\log \frac{1}{\lambda} \right)^2 \mathbb{P}\left\{ \left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \mathbf{1}_{\lambda \leqslant \frac{q}{p} \leqslant 1} \right\},
\end{aligned}$$

where

- the first inequality holds because for any $a, b \in \mathbb{R}$, $(a+b)^2 \leqslant 2a^2 + 2b^2$,
- the second inequality holds because r is decreasing and since $0 \leqslant \frac{q}{p} \leqslant 1$, $\left(\frac{q^{1/2}}{p^{1/2}} - 1 \right)^2 \leqslant 1$,
- the third inequality holds because $r(x) \leqslant 2\log(1/x)$ for $x \in (0, 1/2]$.

All in all,

$$\mathbb{P}\left\{\left(\log\frac{p}{q}\right)^2\mathbf{1}_{\lambda\leqslant\frac{q}{p}\leqslant\frac{1}{\lambda}}\right\}\leqslant 8\left(1+\left(\log\frac{1}{\lambda}\right)^2\right)\mathbb{P}\left\{\left(\frac{q^{1/2}}{p^{1/2}}-1\right)^2\mathbf{1}_{\lambda\leqslant\frac{q}{p}\leqslant\frac{1}{\lambda}}\right\}.$$

□

C.4. Proof of Theorem 15. The proof follows the same lines as the proof of Theorem 5 of [4]. The main difference is that their Proposition 8 is changed into the following one:

PROPOSITION 17. *Let $(Y_t)_{t\in S}$ be a family of real-valued random variables, indexed by a countable subset S of a linear space E of finite dimension D. Let R_2 and R_∞ be two nonnegative random functions on $S\times S$ that almost surely satisfy the triangle inequality. Assume that, for all $s,u\in S$, the increments Y_s-Y_u satisfy (26) with respect to R_2 and R_∞. Assume that there exists two deterministic seminorms N_2 and N_∞ on E, a positive constant β and nonnegative constants v,w satisfying (27). Fix $t_0\in S$. For $\sigma,\sigma'\geqslant 0$, let*

$$\mathcal{B}(\sigma,\sigma')=\left\{s\in S\ :\ R_2(s,t_0)\leqslant\sigma \text{ and } R_\infty(s,t_0)\leqslant\sigma'\right\}.$$

Then, there exists a numerical constant $\kappa>0$ (for instance $\kappa=30$) such that for all $x\geqslant 0$ and $\sigma,\sigma'>0$,

$$\mathbb{P}\left(\sup_{t\in\mathcal{B}(\sigma,\sigma')}(Y_t-Y_{t_0})\geqslant\kappa\left[(\sigma\wedge v^\beta)\sqrt{x+D+\frac{D}{\beta}\log\left(\frac{v^\beta}{\sigma}\vee\frac{w^\beta}{\sigma'}\vee e\right)}\right.\right.$$
$$\left.\left.+(\sigma'\wedge w^\beta)\left(x+D+\frac{D}{\beta}\log\left(\frac{v^\beta}{\sigma}\vee\frac{w^\beta}{\sigma'}\vee e\right)\right)\right]\right)\leqslant e^{-x}.$$

Following this, the same peeling argument is enough to conclude. Let us now prove this proposition.

Without loss of generality, we may assume that S is finite (see [4]). Note that $\mathcal{B}(\sigma,\sigma')=\mathcal{B}(\sigma\wedge v^\beta,\sigma'\wedge w^\beta)$ by (27), so, in what follows, we assume $\sigma\leqslant v^\beta$ and $\sigma'\leqslant w^\beta$.

LEMMA 18. *There exists a sequence of finite partitions $(\mathcal{A}_k)_{k\in\mathbb{N}}$ of S satisfying $\mathcal{A}_0=\{S\}$ and*

$$\begin{cases}\forall k\in\mathbb{N},\ \mathcal{A}_{k+1}\subset\mathcal{A}_k \text{ in the sense that: } \forall B\in\mathcal{A}_{k+1},\exists C\in\mathcal{A}_k \text{ s.t. } B\subset C,\\ \forall k\geqslant 1,\ \forall B\in\mathcal{A}_k,\ \forall s,u\in B,\ N_2(s-u)\leqslant 2^{-k/\beta}\sigma^{1/\beta}\quad\text{and}\quad N_\infty(s-u)\leqslant 2^{-k/\beta}(\sigma')^{1/\beta},\\ \forall k\geqslant 1,\ |\mathcal{A}_k|\leqslant\left(\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)^D 4^{(3+1/\beta)Dk}.\end{cases}$$

PROOF OF LEMMA 18. Let us recall a result from [14, Lemma 4.5] also used by [7]. This result is originally formulated for norms, but extends naturally to seminorms by applying it to the quotient of E by the kernel of the seminorm.

LEMMA 19. *Let N be an arbitrary seminorm on S and $\mathcal{B}_N(0,1)$ its corresponding unit ball. For all $\delta\in(0,1]$, the minimum number of balls of radius δ which is necessary to cover $\mathcal{B}_N(0,1)$ is at most $(1+2\delta^{-1})^D$.*

In the following proof, we build separately for each seminorm N_j a sequence of partitions $(\mathcal{A}_{j,k})_{k\in\mathcal{N}}$ with $j\in\{2,+\infty\}$. The sequence of partitions of Lemma 18 is then obtained by choosing, for $k\geqslant 0$, the partition $\mathcal{A}_k$ defined by

$$\mathcal{A}_k=\left\{A_2\cap A_\infty, A_2\in\mathcal{A}_{2,k}, A_\infty\in\mathcal{A}_{\infty,k}\right\}.$$

For the seminorm N_2. let $\mathcal{A}_{2,0} = S$. By (27), $S \subset \mathcal{B}_{N_2}(0, v)$. Applying Lemma 19 to the norm $v^{-1}N_2$ and $\delta = 2^{-1-1/\beta}v^{-1}\sigma^{1/\beta}$ means that the minimum number of balls of radius $2^{-1-1/\beta}\sigma^{1/\beta}$ which are necessary to cover $\mathcal{B}_{N_2}(0, v)$ is upper bounded by $\left(1 + 2^{2+1/\beta}v\sigma^{-1/\beta}\right)^D \leqslant \left(2^{3+1/\beta}v\sigma^{-1/\beta}\right)^D$. Let $(B_1, \dots, B_p)$ be such a minimal covering. Let $C_1 = B_1$ and for $j \in \{2, \dots, p\}$, define the set C_j as

$$C_j = B_j \backslash \bigcup_{i<j} B_i.$$

The sequence $\mathcal{A}_{2,1} = (C_j)_{j\in\{1,\dots,p\}}$ forms a partition of S, each set of which has a diameter at most $2^{-1/\beta}\sigma^{1/\beta}$.

For $k \geqslant 1$, proceed by induction using Lemma 19. Assume that there exists a partition $\mathcal{A}_{2,k}$ such that $|\mathcal{A}_{2,k}| \leqslant (2^{3+1/\beta}v\sigma^{-1/\beta})^D \cdot 2^{(2+1/\beta)D(k-1)}$ and such that each element of $\mathcal{A}_{2,k}$ is a subset of a ball of radius $2^{-k/\beta-1}\sigma^{1/\beta}$ for the norm N_2. By applying Lemma 19 to $2^{k/\beta+1}\sigma^{-1/\beta}N_2$ and $\delta = 2^{-1/\beta}$, each element of $\mathcal{A}_{2,k}$ can be covered by at most $(1 + 2^{1+1/\beta})^D \leqslant 2^{(2+1/\beta)D}$ balls of radius $2^{-(k+1)/\beta-1}\sigma^{1/\beta}$, and therefore be partitioned into at most $2^{(2+1/\beta)D}$ sets of diameter $2^{-(k+1)/\beta}\sigma^{1/\beta}$, each contained in a ball of radius $2^{-(k+1)/\beta-1}\sigma^{1/\beta}$. $\mathcal{A}_{2,k+1}$ is therefore a partition containing at most $(2^{3+1/\beta}v\sigma^{-1/\beta})^D \cdot 2^{(2+1/\beta)Dk}$ elements.

For the seminorm N_∞. the reasoning is the same and produces a partition $\mathcal{A}_{\infty,k+1}$ containing at most $(2^{3+1/\beta}w(\sigma')^{-1/\beta})^D \cdot 2^{(2+1/\beta)Dk}$ element for all $k \geqslant 0$.

Final partition. By construction, for all $k \in \mathbb{N}$, $\mathcal{A}_{k+1} \subset \mathcal{A}_k$. Moreover, for all $k \geqslant 1$, $A \in \mathcal{A}_k$ and $s, u \in A$, $N_2(s-u) \leqslant 2^{-k/\beta}\sigma^{1/\beta}$ and $N_\infty(s-u) \leqslant 2^{-k/\beta}(\sigma')^{1/\beta}$, and finally,

$$\begin{aligned} |\mathcal{A}_k| &\leqslant |\mathcal{A}_{2,k}| \cdot |\mathcal{A}_{\infty,k}| \leqslant (2^{3+1/\beta}v\sigma^{-1/\beta})^D(2^{3+1/\beta}w(\sigma')^{-1/\beta})^D \cdot 4^{(2+1/\beta)D(k-1)} \\ &\leqslant \left(\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)^D 4^{(3+1/\beta)Dk} \end{aligned}$$

□

Let $(\mathcal{A}_k)_{k\geqslant 0}$ be a sequence of partitions as in Lemma 18. For all $k \in \mathbb{N}^*$ and all $A \in \mathcal{A}_k$, pick a (deterministic) element $t_k(A) \in A$. For any $t \in S$ and $k \geqslant 1$, there exists a unique $A \in \mathcal{A}_k$ such that $t \in A$. Let $\pi_k(t) = t_k(A)$. Let also $\pi_0(t) = t_0$. Since S is finite, the following decomposition holds and contains a finite number of non-zero terms:

$$Y_t - Y_{t_0} = \sum_{k\geqslant 0} \left(Y_{\pi_{k+1}(t)} - Y_{\pi_k(t)}\right).$$

For $k \geqslant 0$, let $E_k = \{(\pi_k(t), \pi_{k+1}(t)), t \in S\}$ and for $k \geqslant 1$, let

$$\begin{cases} z_0 = \dfrac{3}{2}\sigma\sqrt{2(x + \log(2|E_0|))} + \dfrac{3}{2}\sigma'(x + \log(2|E_0|)), \\ z_k = 2^{-k}\left(\sigma\sqrt{2(x + \log(2^{k+1}|E_k|))} + \sigma'(x + \log(2^{k+1}|E_k|))\right). \end{cases}$$

Let

$$\begin{aligned} H &= \frac{3}{2}\sigma\sqrt{2\log(2|E_0|)} + \frac{3}{2}\sigma'\log(2|E_0|) + \sum_{k\geqslant 1} 2^{-k}\left(\sigma\sqrt{2\log(2^{k+1}|E_k|)} + \sigma'\log(2^{k+1}|E_k|)\right) \\ &= \frac{1}{2}\sigma\sqrt{2\log(2|E_0|)} + \frac{1}{2}\sigma'\log(2|E_0|) + \sum_{k\geqslant 0} 2^{-k}\left(\sigma\sqrt{2\log(2^{k+1}|E_k|)} + \sigma'\log(2^{k+1}|E_k|)\right). \end{aligned}$$

Finally, let

$$z = H + \frac{5}{2}\sigma\sqrt{2x} + \frac{5}{2}\sigma' x, \tag{33}$$

so that $z \geqslant \sum_{k\geqslant 0} z_k$. By definition,

$$\mathbb{P}\left(\sup_{t\in\mathcal{B}(\sigma,\sigma')}(Y_t - Y_{t_0}) \geqslant z\right) \leqslant \mathbb{P}\left(\exists t \in \mathcal{B}(\sigma,\sigma'), \exists k \geqslant 0, Y_{\pi_{k+1}(t)} - Y_{\pi_k(t)} \geqslant z_k\right)$$

$$\leqslant \mathbb{P}\left(\sup_{t\in\mathcal{B}(\sigma,\sigma')}\left(Y_{\pi_1(t)} - Y_{t_0}\right) \geqslant z_0\right) + \sum_{k\geqslant 1}\mathbb{P}\left(\sup_{t\in S}\left(Y_{\pi_{k+1}(t)} - Y_{\pi_k(t)}\right) \geqslant z_k\right).$$

The first term must be handled carefully, since even if $t \in \mathcal{B}(\sigma,\sigma')$, there is no guarantee that $\pi_1(t) \in \mathcal{B}(\sigma,\sigma')$. However, if t is in $\mathcal{B}(\sigma,\sigma')$, since $\pi_1(t)$ and t are in the same element of $\mathcal{A}_1$, by the triangle inequality,

$$R_2(\pi_1(t), t_0) \leqslant R_2(\pi_1(t), t) + R_2(t, t_0) \leqslant \frac{3}{2}\sigma,$$

and likewise $R_\infty(\pi_1(t), t_0) \leqslant \frac{3}{2}\sigma'$. Therefore,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in\mathcal{B}(\sigma,\sigma')}(Y_t - Y_{t_0}) \geqslant z\right) &\leqslant \mathbb{P}\left(\sup_{u \text{ s.t. } (t_0,u)\in E_0 \text{ and } u\in\mathcal{B}(3\sigma/2,3\sigma'/2)}(Y_u - Y_{t_0}) \geqslant z_0\right) \\
&\quad + \sum_{k\geqslant 1}\mathbb{P}\left(\sup_{(s,u)\in E_k}(Y_u - Y_s) \geqslant z_k\right) \\
&\leqslant \sum_{u \text{ s.t. } (t_0,u)\in E_0}\mathbb{P}\left((Y_u - Y_{t_0}) \geqslant z_0 \text{ and } u \in \mathcal{B}(3\sigma/2, 3\sigma'/2)\right) \\
&\quad + \sum_{k\geqslant 1}\mathbb{P}\left(\sup_{(s,u)\in E_k}(Y_u - Y_s) \geqslant z_k\right) \\
&\leqslant \sum_{u \text{ s.t. } (t_0,u)\in E_0}\mathbb{P}\left((Y_u - Y_{t_0}) \geqslant z_0 \text{ and } u \in \mathcal{B}(3\sigma/2, 3\sigma'/2)\right) \\
&\quad + \sum_{k\geqslant 1}\sum_{(s,u)\in E_k}\mathbb{P}\left(Y_u - Y_s \geqslant z_k\right).
\end{aligned}$$

For $k = 0$. Using (26) and the definition of z_0,

$$\begin{aligned}
&\mathbb{P}\left(Y_u - Y_{t_0} \geqslant z_0 \text{ and } u \in \mathcal{B}(3\sigma/2, 3\sigma'/2)\right) \\
&\quad = \mathbb{P}\Bigg(Y_u - Y_{t_0} \geqslant \frac{3}{2}\sigma\sqrt{2(x + \log(2|E_0|))} + \frac{3}{2}\sigma'(x + \log(2|E_0|)) \\
&\qquad\qquad \text{and } R_2(u, t_0) \leqslant \frac{3}{2}\sigma,\ R_\infty(u, t_0) \leqslant \frac{3}{2}\sigma'\Bigg) \leqslant \frac{1}{2|E_0|}e^{-x}.
\end{aligned}$$

For $k \geqslant 1$. Since $\mathcal{A}_{k+1} \subset \mathcal{A}_k$, $\pi_k(t)$ and $\pi_{k+1}(t)$ belong to the same set in $\mathcal{A}_k$. Therefore, for all $(s,u) \in E_k$, $N_2(s-u) \leqslant 2^{-k/\beta}\sigma^{1/\beta}$ and $N_\infty(s-u) \leqslant 2^{-k/\beta}(\sigma')^{1/\beta}$. By assumption,

$R_2(s,u) \leqslant N_2(s-u)^\beta$ and $R_\infty(s,u) \leqslant N_\infty(s-u)^\beta$. Thus, for $(s,u) \in E_k$, $R_2(s,u) \leqslant 2^{-k}\sigma$ and $R_\infty(s,u) \leqslant 2^{-k}\sigma'$ almost surely. By definition of z_k and (26), for all $(s,u) \in E_k$,

$$\mathbb{P}(Y_u - Y_s \geqslant z_k) \leqslant 2^{-(k+1)}|E_k|^{-1}e^{-x}.$$

Summing on all $(s,u) \in E_k$ and all $k \geqslant 0$ leads to

$$\mathbb{P}\left(\sup_{t\in\mathcal{B}(\sigma,\sigma')}(Y_t - Y_{t_0}) \geqslant z\right) \leqslant e^{-x}.$$

It remains to compute H in (33). By construction of $\mathcal{A}_k$, the choice of $\pi_{k+1}(t)$ entirely determines the choice of $\pi_k(t)$. Therefore, $|E_k| \leqslant |\mathcal{A}_{k+1}|$, that is

$$2^{k+1}|E_k| \leqslant 2^{k+1}\left(\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)^D 4^{(3+1/\beta)D(k+1)} \leqslant \left(4^{(4+1/\beta)(k+1)}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)^D$$

and thus

$$H \leqslant \underbrace{\frac{1}{2}\sigma\sqrt{2D\log\left(4^{4+1/\beta}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)} + \frac{1}{2}\sigma' D\log\left(4^{4+1/\beta}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)}_{E}$$
$$+\underbrace{\sum_{k\geqslant 0}2^{-k}\sigma\sqrt{2D\log\left(4^{(4+1/\beta)(k+1)}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)}}_{F} + \underbrace{\sum_{k\geqslant 0}2^{-k}\sigma' D\log\left(4^{(4+1/\beta)(k+1)}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)}_{G}.$$

Let us calculate the three terms separately. Firstly,

$$E \leqslant \frac{\sigma}{2}\sqrt{2D\left((4+\frac{1}{\beta})\log 4 + \frac{1}{\beta}\log(\frac{(vw)^\beta}{\sigma\sigma'})\right)} + \frac{\sigma'}{2}D\left((4+\frac{1}{\beta})\log 4 + \frac{1}{\beta}\log(\frac{(vw)^\beta}{\sigma\sigma'})\right)$$
$$\leqslant \frac{\sigma}{2}\sqrt{2D\left(6+\frac{1}{\beta}\log(\frac{4(vw)^\beta}{\sigma\sigma'})\right)} + \frac{\sigma'}{2}D\left(6+\frac{1}{\beta}\log(\frac{4(vw)^\beta}{\sigma\sigma'})\right).$$

Secondly,

$$G = \sigma' D\sum_{k\geqslant 0}2^{-k}\left((k+1)(4+\frac{1}{\beta})\log 4 + \frac{1}{\beta}\log(\frac{(vw)^\beta}{\sigma\sigma'})\right)$$
$$= 2\sigma' D\left(2(4+\frac{1}{\beta})\log 4 + \frac{1}{\beta}\log(\frac{(vw)^\beta}{\sigma\sigma'})\right)$$
$$\leqslant 2\sigma' D\left(12+\frac{1}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})\right).$$

Thirdly, by concavity of $x \mapsto \sqrt{x}$ and Jensen's inequality,

$$F \leqslant 2\sigma\sqrt{\sum_{k\geqslant 0}2^{-k}D\log\left(4^{(4+1/\beta)(k+1)}\frac{vw}{(\sigma\sigma')^{1/\beta}}\right)} \leqslant 2\sigma\sqrt{2D\left(12+\frac{1}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})\right)}.$$

Thus,

$$H \leqslant \frac{5}{2}\sigma\sqrt{2D\left(12+\frac{1}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})\right)} + \frac{5}{2}\sigma' D\left(12+\frac{1}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})\right).$$

Finally, using the concavity of $x \mapsto \sqrt{x}$ again,

$$z = H + \frac{5}{2}\sigma\sqrt{2x} + \frac{5}{2}\sigma' x \leqslant 5\sigma\sqrt{x + 12D + \frac{D}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})} + \frac{5}{2}\sigma'\left(x + 12D + \frac{D}{\beta}\log(\frac{16(vw)^\beta}{\sigma\sigma'})\right)$$
$$\leqslant 30\left[\sigma\sqrt{x + D + \frac{D}{\beta}\log\left(\frac{v^\beta}{\sigma} \vee \frac{w^\beta}{\sigma'} \vee e\right)} + \sigma'\left(x + D + \frac{D}{\beta}\log\left(\frac{v^\beta}{\sigma} \vee \frac{w^\beta}{\sigma'} \vee e\right)\right)\right].$$

This concludes the proof of Proposition 17.

APPENDIX D: PROOF OF SECTION 4.2

PROOF OF LEMMA 6. Let $m \in \mathcal{M}$. For any $t > k \geqslant 0$, any probability distribution λ on $[h_m]$ and any $\theta \in \Theta_m$, let

$$L_{t,k,\lambda}(\theta) = \log \mathbb{E}_{S_{t-k}\sim\lambda}\left[p_\theta^m(X_t | X_{t-k}^{t-1}, S_{t-k})\right],$$

where p_θ^m is the conditional distribution according to the parameters of the HMM model m. This can be seen as the log of the conditional density of a HMM initialized at time $t-k$ with initial distribution λ. For $k = 0$, we use the convention $p_\theta^m(X_t | X_{t-0}^{t-1}, S_{t-0}) = p_\theta^m(X_t | S_t)$.

Let $\rho = 1 - (C_Q)^{-2}$. Recall that we assumed $\pi_\theta^m, Q_\theta^m \in [(C_Q)^{-1}h_m^{-1}, C_Q h_m^{-1}]$ for all $\theta \in \Theta_m$. As a consequence, by Lemma 3 of [29], for all t, k, k', λ and λ', almost surely,

$$\sup_{\theta\in\Theta_m} |L_{t,k,\lambda}(\theta) - L_{t,k',\lambda'}(\theta)| \leqslant \rho^{k\wedge k'-1}/(1-\rho).$$

In particular, for all $t > k \geqslant 0$, almost surely,

$$\sup_{\theta\in\Theta_m} |\log p_{\theta,t}^m(X_t) - L_{t,k,\pi_\theta^m}(\theta)| \leqslant \rho^{k-1}/(1-\rho).$$

Let $t > k \geqslant 0$. For all $\theta \in \Theta_m$, $x_{t-k}^t \in \mathcal{X}^{k+1}$ and $i \in [h_m]$, let $f_{\theta,i}$ be defined as

$$f_{\theta,i}(x_{t-k}^t) = p_\theta^m(S_k = i, X_t = x_t | X_{t-k}^{t-1} = x_{t-k}^{t-1}, S_{t-k} \sim \pi_\theta^m).$$

Note that $L_{t,k,\pi_\theta^m}(\theta) = \log \sum_{i\in[h_m]} f_{\theta,i}(X_{t-k}^t)$, so that for any $\theta, \delta \in \Theta_m$,

$$|\log p_{\theta,t}^m(X_t) - \log p_{\delta,t}^m(X_t)| \leqslant \left|\log \sum_{i\in[h_m]} (f_{\theta,i}(X_{t-k}^t) - f_{\delta,i}(X_{t-k}^t))\right| + 2\rho^{k-1}/(1-\rho).$$

Note that if $t = k+1$, a simpler equality holds:

$$|\log p_{\theta,t}^m(X_t) - \log p_{\delta,t}^m(X_t)| = \left|\log \sum_{i\in[h_m]} (f_{\theta,i}(X_{t-k}^t) - f_{\delta,i}(X_{t-k}^t))\right|.$$

LEMMA 20. *Assume* $\|\theta - \delta\|_m \leqslant \log 2$, *then for all* $x_0^k \in \mathcal{X}^{k+1}$,

$$\left|\log \sum_{i\in[h_m]} (f_{\theta,i}(x_0^k) - f_{\delta,i}(x_0^k))\right| \leqslant 7(2C_Q^2)^{k+3}\|\theta - \delta\|_m.$$

To simplify this formula, there exists two constants $C_1, C_2 > 1$ such that

$$\left|\log \sum_{i\in[h_m]} (f_{\theta,i}(x_0^k) - f_{\delta,i}(x_0^k))\right| \leqslant \|\theta - \delta\|_m C_1 C_2^k.$$

Therefore, up to changing C_1, almost surely,

$$|\log p^m_{\theta,t}(X_t) - \log p^m_{\delta,t}(X_t)| \leqslant C_1 \min\left(\inf_{k\in\mathbb{N}:0\leqslant k<t} C_2^k \|\theta-\delta\|_m + \rho^k, C_2^{t-1}\|\theta-\delta\|_m\right).$$

A possible choice is $C_2 = 2C_Q^2$ and $C_1 = 56C_Q^6$.

Solving $C_2^k\|\theta-\delta\|_m = \rho^k$ results in the trade-off $k^\star = \lceil \log(1/\|\theta-\delta\|_m)/\log(C_2/\rho)\rceil$, which ensures that as long as $\|\theta-\delta\|_m \leqslant \log 2$, almost surely,

$$\begin{aligned}|\log p^m_{\theta,t}(X_t) - \log p^m_{\delta,t}(X_t)| &\leqslant 2C_1 C_2^{k^\star}\|\theta-\delta\|_m \\ &\leqslant 2C_1C_2\|\theta-\delta\|_m^{1-\frac{\log C_2}{\log(C_2/\rho)}},\end{aligned}$$

so we may take

$$\begin{aligned}\beta_m^{-1} = 1 - \frac{\log C_2}{\log(C_2/\rho)} = \frac{\log\rho^{-1}}{\log C_2 + \log\rho^{-1}} &= \frac{-\log(1-(C_Q)^{-2})}{\log(2C_Q^2) - \log(1-(C_Q)^{-2})} \\ &\sim_{C_Q\to+\infty} \frac{C_Q^{-2}}{2\log C_Q}\end{aligned}$$

and since $L_m \geqslant (2C_1C_2)$, it proves that Assumption 2 holds when $\|\theta-\delta\|_m \leqslant \log 2$. When $\|\theta-\delta\|_m \geqslant \log 2$, we always have

$$|\log p^m_{\theta,t}(X_t) - \log p^m_{\delta,t}(X_t)| \leqslant 2\log C_m^g \leqslant 3\log C_m^g \|\theta-\delta\|_m^{\beta_m},$$

and thus Assumption 2 holds. □

D.1. Proof of Lemma 20. Let $x_0^k \in \mathcal{X}^{k+1}$. Using the approach of Appendix A of [25], one can write $f_{\theta,i}$ as the product of matrices

$$f_{\theta,i}(x_0^k) = \left(\mu^\theta_{0|k-1}F^\theta_{1|k-1}\dots F^\theta_{k-1|k-1}Q^m_\theta\right)_i \nu^m_{\theta,i}(x_k) \tag{34}$$

where

$$\beta^\theta_{u|k}(s_u) = \sum_{s_{u+1}^k\in[h_m]^{k-u}} Q^m_\theta(s_u,s_{u+1})\nu^m_{\theta,s_{u+1}}(x_{u+1})\dots Q^m_\theta(s_{k-1},s_k)\nu^m_{\theta,s_k}(x_k),$$

for $0\leqslant u\leqslant k-1$ and $\beta^\theta_{k|k}(i) = 1$ for all $i\in[h_m]$,

$$\mu^\theta_{0|k}(i) = \frac{\pi^m_\theta(i)\beta^\theta_{0|k}(i)\nu^m_{\theta,i}(x_0)}{\sum_{j\in[h_m]}\pi^m_\theta(j)\beta^\theta_{0|k}(j)\nu^m_{\theta,j}(x_0)}$$

$$\text{and}\quad F^\theta_{u|k}(s_{u-1},s_u) = \frac{\beta^\theta_{u|k}(s_u)Q^m_\theta(s_{u-1},s_u)\nu^m_{\theta,s_u}(x_u)}{\sum_{i\in[h_m]}\beta^\theta_{u|k}(i)Q^m_\theta(s_{u-1},i)\nu^m_{\theta,i}(x_u)}.$$

An intuition of these quantities, and a way to check (34), is given by the identities

$$\begin{aligned}\beta^\theta_{u|k}(s_u) &= p^m_\theta(X^{k+1}_{u+2} = x^k_{u+1}|S_{u+1}=s_u),\\ \mu^\theta_{0|k}(i) &= p^m_\theta(S_1 = i|X_1^{k+1} = x_0^k),\\ F^\theta_{u|k}(s_{u-1},s_u) &= p^m_\theta(S_{u+1}=s_u|X^{k+1}_{u+1} = x^k_u, S_u = s_{u-1}),\end{aligned}$$

(keeping in mind that under p^m_θ, $S_1\sim\pi^m_\theta$), so that

$$\left(\mu^\theta_{0|k}F^\theta_{1|k}\dots F^\theta_{k|k}\right)_i = p^m_\theta(S_{k+1}=i|X_1^{k+1}=x_0^k).$$

The following Lemmas follow the proofs of Appendix B.2.2 of [40].

LEMMA 21. *Assume* $\|\theta-\delta\|_m \leqslant \log 2$, *then*

$$\sup_{0\leqslant u\leqslant k} \frac{\sum_{i\in[h_m]} |\beta^\theta_{u|k}(i)\nu^m_{\theta,i}(x_u) - \beta^\delta_{u|k}(i)\nu^m_{\delta,i}(x_u)|}{\sum_{i\in[h_m]} \beta^\theta_{u|k}(i)\nu^m_{\theta,i}(x_u)} \leqslant \|\theta-\delta\|_m (C_Q^{-1} + (\log 2)^{-1})(2C_Q^2)^{k+1}.$$

PROOF. Using minimalist notations,

$$\begin{aligned}
&\sum_{i\in[h_m]} |\beta^\theta_{u|k}(i)\nu^m_{\theta,i}(x_u) - \beta^\delta_{u|k}(i)\nu^m_{\delta,i}(x_u)| \\
&\quad = \sum_{s_u\in[h_m]} \Bigg| \Bigg(\sum_{s^k_{u+1}\in[h_m]^{k-u}} Q^\theta_{u,u+1}\nu^\theta_{u+1}Q^\theta_{u+1,u+2}\dots\nu^\theta_k\Bigg)\nu^\theta_u \\
&\qquad\quad - \Bigg(\sum_{s^k_{u+1}\in[h_m]^{k-u}} Q^\delta_{u,u+1}\nu^\delta_{u+1}Q^\delta_{u+1,u+2}\dots\nu^\delta_k\Bigg)\nu^\delta_u \Bigg| \\
&\quad \leqslant \sum_{j=u+1}^k \sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\nu^\theta_{u+1}\dots\nu^\theta_{j-1}|Q^\theta_{j-1,j} - Q^\delta_{j-1,j}|\nu^\delta_j\dots Q^\delta_{k-1,k}\nu^\delta_k \\
&\qquad + \sum_{j=u}^k \sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\nu^\theta_{u+1}\dots Q^\theta_{j-1,j}|\nu^\theta_j - \nu^\delta_j|Q^\delta_{j,j+1}\dots Q^\delta_{k-1,k}\nu^\delta_k.
\end{aligned}$$

Keep in mind in what follows that Assumption 4-HMM and the definition of $\|\cdot\|_m$ entail, for all j,

$$|\nu^\theta_j - \nu^\delta_j| \leqslant \nu^\theta_j(e^{\|\theta-\delta\|_m} - 1) \qquad \text{and} \qquad \nu^\delta_j \leqslant \nu^\theta_j e^{\|\theta-\delta\|_m}. \tag{35}$$

By definition of $\|\cdot\|_m$, for all $j\in\{u+1,\dots,k\}$,

$$\begin{aligned}
&\sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\dots Q^\theta_{j-2,j-1}\nu^\theta_{j-1}|Q^\theta_{j-1,j} - Q^\delta_{j-1,j}|\nu^\delta_j\dots Q^\delta_{k-1,k}\nu^\delta_k \\
&\qquad \leqslant \|\theta-\delta\|_m h_m^{-1}(C_Q h_m^{-1})^{k-j} \sum_{s^{j-1}_u\in[h_m]^{j-u}} \nu^\theta_u Q^\theta_{u,u+1}\dots Q^\theta_{j-2,j-1}\nu^\theta_{j-1} \sum_{s_j\in[h_m]} \nu^\delta_j \cdots \sum_{s_k\in[h_m]} \nu^\delta_k
\end{aligned}$$

and for all $j\in\{u,\dots,k\}$

$$\begin{aligned}
\sum_{i\in[h_m]} \beta^\theta_{u|k}(i)\nu^m_{\theta,i}(x_u) &= \sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\dots Q^\theta_{j-2,j-1}\nu^\theta_{j-1}Q^\theta_{j-1,j}\nu^\theta_j\dots Q^\theta_{k-1,k}\nu^\theta_k \\
&\geqslant (C_Q h_m)^{-(k-j+1)} \sum_{s^{j-1}_u\in[h_m]^{j-u}} \nu^\theta_u Q^\theta_{u,u+1}\dots Q^\theta_{j-2,j-1}\nu^\theta_{j-1} \sum_{s_j\in[h_m]} \nu^\theta_j \cdots \sum_{s_k\in[h_m]} \nu^\theta_k,
\end{aligned}$$

so that

$$\begin{aligned}
&\frac{\sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\dots\nu^\theta_{j-1}|Q^\theta_{j-1,j} - Q^\delta_{j-1,j}|\nu^\delta_j\dots\nu^\delta_k}{\sum_{s^k_u\in[h_m]^{k-u+1}} \nu^\theta_u Q^\theta_{u,u+1}\dots\nu^\theta_{j-1}Q^\theta_{j-1,j}\nu^\theta_j\dots\nu^\theta_k} \\
&\qquad\qquad \leqslant \|\theta-\delta\|_m e^{(k-j+1)\|\theta-\delta\|_m}(C_Q)^{2(k-j)+1}.
\end{aligned}$$

Likewise, for all $j \in \{u, \dots, k\}$,

$$\frac{\sum_{s_u^k \in [h_m]^{k-u+1}} \nu_u^\theta Q_{u,u+1}^\theta \dots Q_{j-1,j}^\theta |\nu_j^\theta - \nu_j^\delta| Q_{j,j+1}^\delta \dots Q_{k-1,k}^\delta \nu_k^\delta}{\sum_{s_u^k \in [h_m]^{k-u+1}} \nu_u^\theta Q_{u,u+1}^\theta \dots Q_{j-1,j}^\theta \nu_j^\theta Q_{j,j+1}^\theta \dots Q_{k-1,k}^\theta \nu_k^\theta} \leqslant e^{(k-j)\|\theta-\delta\|_m}(e^{\|\theta-\delta\|_m} - 1) C_Q^{2(k-j)}.$$

Therefore,

$$\begin{aligned}
&\frac{\sum_{i \in [h_m]} |\beta_{u|k}^\theta(i)\nu_{\theta,i}^m(x_u) - \beta_{u|k}^\delta(i)\nu_{\delta,i}^m(x_u)|}{\sum_{i \in [h_m]} \beta_{u|k}^\theta(i)\nu_{\theta,i}^m(x_u)} \\
&\leqslant \|\theta-\delta\|_m C_Q^{-1} \sum_{j=u}^{k-1} e^{(k-j)\|\theta-\delta\|_m} (C_Q)^{2(k-j)} + (e^{\|\theta-\delta\|_m} - 1) \sum_{j=u}^{k} e^{(k-j)\|\theta-\delta\|_m} C_Q^{2(k-j)} \\
&\leqslant \left(\|\theta-\delta\|_m C_Q^{-1} + (e^{\|\theta-\delta\|_m} - 1)\right) \sum_{j=u}^{k} (e^{\|\theta-\delta\|_m} C_Q^2)^{k-j} \\
&\leqslant \|\theta-\delta\|_m (C_Q^{-1} + (\log 2)^{-1}) \frac{\left(2C_Q^2\right)^{k-u+1} - 1}{2C_Q^2 - 1} \qquad \text{when } \|\theta-\delta\|_m \leqslant \log 2 \\
&\leqslant \|\theta-\delta\|_m (C_Q^{-1} + (\log 2)^{-1}) (2C_Q^2)^{k-u+1} \qquad \text{since } C_Q \geqslant 1.
\end{aligned}$$

□

LEMMA 22. *Assume* $\|\theta-\delta\|_m \leqslant \log 2$*, then*

$$\|\mu_{0|k}^\theta - \mu_{0|k}^\delta\|_1 \leqslant 3(2C_Q^2)^{k+2} \|\theta-\delta\|_m$$

and

$$\sup_{0 \leqslant u \leqslant k} \sup_{i \in [h_m]} \|F_{u|k}^\theta(i,\cdot) - F_{u|k}^\delta(i,\cdot)\|_1 \leqslant 3(2C_Q^2)^{k+2} \|\theta-\delta\|_m.$$

PROOF. With minimalist notations,

$$\begin{aligned}
\sum_i |\mu_i^\theta - \mu_i^\delta| &= \sum_i \left| \frac{\pi_i^\theta \beta_i^\theta \nu_i^\theta}{\sum_j \pi_j^\theta \beta_j^\theta \nu_j^\theta} - \frac{\pi_i^\delta \beta_i^\delta \nu_i^\delta}{\sum_j \pi_j^\delta \beta_j^\delta \nu_j^\delta} \right| \\
&\leqslant \frac{\sum |\pi^\theta \beta^\theta \nu^\theta - \pi^\delta \beta^\delta \nu^\delta|}{\sum \pi^\theta \beta^\theta \nu^\theta} + \left| \frac{1}{\sum \pi^\theta \beta^\theta \nu^\theta} - \frac{1}{\sum \pi^\delta \beta^\delta \nu^\delta} \right| \sum \pi^\delta \beta^\delta \nu^\delta \\
&\leqslant 2 \frac{\sum |\pi^\theta \beta^\theta \nu^\theta - \pi^\delta \beta^\delta \nu^\delta|}{\sum \pi^\theta \beta^\theta \nu^\theta} \\
&\leqslant 2 C_Q h_m \frac{\sum |\pi^\theta \beta^\theta \nu^\theta - \pi^\delta \beta^\delta \nu^\delta|}{\sum \beta^\theta \nu^\theta}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum |\mu^\theta - \mu^\delta| &\leqslant 2 C_Q h_m \left(\frac{\sum |\pi^\theta - \pi^\delta| \beta^\theta \nu^\theta}{\sum \beta^\theta \nu^\theta} + \frac{\sum \pi^\delta |\beta^\theta \nu^\theta - \beta^\delta \nu^\delta|}{\sum \beta^\theta \nu^\theta} \right) \\
&\leqslant 2 C_Q h_m \left(\|\theta-\delta\|_m h_m^{-1} + C_Q h_m^{-1} \frac{\sum |\beta^\theta \nu^\theta - \beta^\delta \nu^\delta|}{\sum \beta^\theta \nu^\theta} \right)
\end{aligned}$$

$$\leqslant \|\theta-\delta\|_m 2C_Q\left(1+C_Q(C_Q^{-1}+(\log 2)^{-1})(2C_Q^2)^{k+1}\right) \quad \text{by Lemma 21}$$
$$\leqslant 3(2C_Q^2)^{k+2}\|\theta-\delta\|_m.$$

The control of $\sum_{j\in[K]}|F^\theta_{u|k}(i,j)-F^\delta_{u|k}(i,j)|$ is exactly the same after replacing π^m_θ by $Q^m_\theta(i,\cdot)$ and likewise for δ. □

We now have the tools to prove Lemma 20. We again use minimalist notations. First,

$$\sum_{i\in[h_m]}|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k})_i-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k})_i| \leqslant \sum_{i\in[h_m]}|((\mu^\theta-\mu^\delta)F^\theta_{1|k}\dots F^\theta_{k|k})_i|$$
$$+\sum_{u=1}^{k}\sum_{i\in[h_m]}|(\mu^\delta F^\delta_{1|k}\dots F^\delta_{u-1|k}(F^\delta_{u|k}-F^\theta_{u|k})F^\theta_{u+1|k}\dots F^\theta_{k|k})_i|.$$

Then, since $F^\theta_{u|k}$ and $F^\delta_{u|k}$ are transition matrices (and thus are 1-Lipschitz linear operators of $\mathbf{L}^1([h_m])$),

$$\|\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}-\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}\|_1 \leqslant \|\mu^\delta-\mu^\theta\|_1+\sum_{u=1}^{k}\sup_{i\in[h_m]}\|F^\theta_{u|k}(i,\cdot)-F^\delta_{u|k}(i,\cdot)\|_1.$$

By Lemma 22, both $\|\mu^\delta-\mu^\theta\|_1$ and $\sup_{1\leqslant u\leqslant k}\sup_{i\in[h_m]}\|F^\theta_{u|k}(i,\cdot)-F^\delta_{u|k}(i,\cdot)\|_1$ are upper bounded by $C_F\|\theta-\delta\|_m$ for $C_F=3(2C_Q^2)^{k+2}$, so that

$$\|\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}-\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}\|_1 \leqslant 2C_F\|\theta-\delta\|_m.$$

Therefore,

$$\frac{\sum_{i\in[h_m]}|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i\nu^\theta_i-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}Q^\delta)_i\nu^\delta_i|}{\sum_{i\in[h_m]}(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i\nu^\theta_i}$$
$$\leqslant \frac{\sum_i(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i|\nu^\theta_i-\nu^\delta_i|}{\sum_{i\in[h_m]}(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i\nu^\theta_i}$$
$$+\frac{\sum_i|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}Q^\delta)_i|\nu^\delta_i}{\sum_{i\in[h_m]}(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i\nu^\theta_i}$$
$$\leqslant \|\theta-\delta\|_m(\log 2)^{-1}+2\sup_i\frac{|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}Q^\delta)_i|}{(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i}$$

by (35) when $\|\theta-\delta\|_m\leqslant\log 2$. Then, for any $i\in[h_m]$, since $(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i\geqslant(C_Qh_m)^{-1}$,

$$\frac{|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k}Q^\delta)_i|}{(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k}Q^\theta)_i}$$
$$\leqslant C_Qh_m\sum_j|(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k})_j-(\mu^\delta F^\delta_{1|k}\dots F^\delta_{k|k})_j|Q^\delta_{j,i}$$
$$+C_Qh_m\sum_j(\mu^\theta F^\theta_{1|k}\dots F^\theta_{k|k})_j|Q^\theta_{j,i}-Q^\delta_{j,i}|$$

$$\leqslant C_Q^2 \sum_j |(\mu^\theta F_{1|k}^\theta \dots F_{k|k}^\theta)_j - (\mu^\delta F_{1|k}^\delta \dots F_{k|k}^\delta)_j|$$

$$+ C_Q \|\theta - \delta\|_m \sum_j (\mu^\theta F_{1|k}^\theta \dots F_{k|k}^\theta)_j$$

$$\leqslant 2C_Q^2 C_F \|\theta - \delta\|_m + C_Q \|\theta - \delta\|_m .$$

As a consequence, by symmetry of the roles of θ and δ,

$$\left| \log \sum_{i \in [h_m]} (f_{\theta,i}(x_0^k) - f_{\delta,i}(x_0^k)) \right| \leqslant \frac{\sum_{i\in[h_m]} |f_{\theta,i}(x_0^k) - f_{\delta,i}(x_0^k)|}{\left(\sum_{i\in[h_m]} f_{\theta,i}(x_0^k)\right) \wedge \left(\sum_{i\in[h_m]} f_{\delta,i}(x_0^k)\right)}$$

$$\leqslant \|\theta - \delta\|_m [12 C_Q^2 (2C_Q^2)^{k+2} + 2C_Q + (\log 2)^{-1}] \leqslant 7(2C_Q^2)^{k+3} \|\theta - \delta\|_m .$$

APPENDIX E: PROOF OF SECTION 4.3

PROOF OF LEMMA 7. For $u \in \mathbb{R}$, let $\phi(u) = e^u - u - 1$. Let $t \in [n]$. By definition,

$$\mathbb{E}\left[\log\left(\frac{p_t^\star(X_t)}{p_t(X_t)}\right) \Big| \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\phi\left(\log\left(\frac{p_t(X_t)}{p_t^\star(X_t)}\right)\right) \Big| \mathcal{F}_{t-1}\right].$$

Let $x = \log\left(\frac{p_t(X_t)}{p_t^\star(X_t)}\right)$. By hypothesis

$$|x| = \log\left(\frac{p_t(X_t)}{p_t^\star(X_t)}\right) \mathbf{1}_{x \geqslant 0} + \log\left(\frac{p_t^\star(X_t)}{p_t(X_t)}\right) \mathbf{1}_{x<0}$$

If $x \geqslant 0$, then $\varepsilon \leqslant p_t^\star(X_t) \leqslant p_t(X_t) \leqslant \varepsilon^{-1}$, so that $\log\left(\frac{p_t(X_t)}{p_t^\star(X_t)}\right) \leqslant \log(\varepsilon^{-2})$. The same inequality holds by symmetry when $x < 0$. Therefore,

$$|x| \leqslant \log(\varepsilon^{-2}).$$

By convexity, note that ϕ is non negative. Let us apply the Taylor Lagrange formula in 0. There exists $c \in]0, u[$ such that

$$\phi(u) = \phi(0) + \phi'(0)u + \phi''(c)\frac{u^2}{2}.$$

Since, $\phi(0) = \phi'(0) = 0$, the previous equation simply reads $\phi(u) = \phi''(c)\frac{u^2}{2}$. Since the function $u \to \phi''(u) = e^u$ is increasing, it holds that for all $u \in [-\log(\varepsilon^{-2}), \log(\varepsilon^{-2})]$

$$\phi''(-\log(\varepsilon^{-2}))\frac{u^2}{2} = \frac{\varepsilon^2}{2}u^2 \leqslant \phi(u) \leqslant \phi''(\log(\varepsilon^{-2}))\frac{u^2}{2} = \frac{1}{2\varepsilon^2}u^2.$$

That is,

$$\frac{\varepsilon^2}{2}\mathbb{E}\left[\log\left(\frac{p_t^\star(X_t)}{p_t(X_t)}\right)^2 \Big| \mathcal{F}_{t-1}\right] \leqslant \mathbb{E}\left[\log\left(\frac{p_t^\star(X_t)}{p_t(X_t)}\right) \Big| \mathcal{F}_{t-1}\right] \leqslant \frac{1}{2\varepsilon^2}\mathbb{E}\left[\log\left(\frac{p_t^\star(X_t)}{p_t(X_t)}\right)^2 \Big| \mathcal{F}_{t-1}\right].$$

□

APPENDIX F: PROOF OF SECTION 4.4

F.1. Proof of Section 4.4.1.

PROOF OF LEMMA 8. In [42, Lemma 7.23], the author shows that for two densities f and g, it holds that

$$2\,\mathrm{H}^2(f,g) \leqslant \mathrm{KL}(f,g)$$

where H is the Hellinger distance defined by,

$$\mathrm{H}^2(f,g) = \frac{1}{2}\int \left(\sqrt{f}-\sqrt{g}\right)^2 .$$

Let us apply this result to $p_t^\star$ and p_t:

$$\sum_{k=1}^{K} p_t^\star(k)\log\frac{p_t^\star(k)}{p_t(k)} \geqslant \frac{1}{2}\sum_{k=1}^{K}\left(\sqrt{p_t^\star(k)}-\sqrt{p_t(k)}\right)^2 \geqslant \frac{1}{8}\sum_{k=1}^{K}\left(p_t^\star(k)-p_t(k)\right)^2 ,$$

where the last inequality holds because, by the Taylor-Lagrange formula, for all $x,y\in[\varepsilon,1]$,

$$|\sqrt{x}-\sqrt{y}| \geqslant \frac{1}{2}|x-y|.$$

Therefore,

$$\mathbf{K}_n(p) \geqslant \frac{1}{8T_\varepsilon}\sum_{t=1}^{T_\varepsilon}\sum_{k=1}^{K}\left(p_t^\star(k)-p_t(k)\right)^2 .$$

□

F.2. Proofs of Section 4.4.2.

F.2.1. *Unbiased estimator.* Let's first verify the following claim.

PROPOSITION 23. *In the procedure described in Section 4.4.2, $\hat{L}_t^m(J)$ is an unbiased estimator of the loss g_J^m with respect to the distribution p_s.*

PROOF. Given the past $\sigma(X_1^{t-1})$, taking the conditional expectation with respect to p_t

$$\begin{aligned}
\mathbb{E}\left[\hat{L}_t^m(J)|X_1^{t-1}\right] &= \frac{g_J^m}{p_t(J)}\mathbb{P}(X_t\in J|X_1^{t-1}) \\
&= \frac{g_J^m}{p_t(J)}\left(\mathbb{P}(X_t\in J, I_t=J|X_1^{t-1}) + \underbrace{\mathbb{P}(X_t\in J, I_t\neq J|X_1^{t-1})}_{=0}\right) \\
&= \frac{g_J^m}{p_t(J)}\mathbb{P}(I_t=J|X_1^{t-1})\underbrace{\mathbb{P}(X_t\in J|I_t=J, X_1^{t-1})}_{=1} \\
&= \frac{g_J^m}{p_t(J)}p_t(J) = g_J^m .
\end{aligned}$$

□

F.2.2. *Proof of Proposition 9.* Equation (8) can be written

$$p^m_{\theta^m,t+1}(I) = \frac{p^m_{\theta^m,t}(I)\exp\left(-\frac{\theta^m_I/\sqrt{T}}{p_{\theta^m,t}(I)}\right)\mathbf{1}_{X_t\in I}}{1-p^m_{\theta^m,t}(I)+p^m_{\theta^m,t}(I)\exp\left(-\frac{\theta^m_I/\sqrt{T}}{p_{\theta^m,t}(I)}\right)} + \sum_{J\neq I}\frac{p^m_{\theta^m,t}(I)\mathbf{1}_{X_t\in J}}{1-p^m_{\theta^m,t}(J)+p^m_{\theta^m,t}(J)\exp\left(-\frac{\theta^m_J/\sqrt{T}}{p_{\theta^m,t}(J)}\right)}.$$

Since for all $q>0$, $1-q+q\exp\left(-\frac{\theta^m_I/\sqrt{T}}{q}\right)\leqslant 1$,

$$p^m_{\theta^m,t+1}(I)\geqslant p^m_{\theta^m,t}(I)\exp\left(-\frac{\theta^m_I/\sqrt{T}}{p_{\theta^m,t}(I)}\right)\mathbf{1}_{X_t\in I}+p^m_{\theta^m,t}(I)\sum_{J\neq I}\mathbf{1}_{X_t\in J},$$

so that

$$p^m_{\theta^m,t+1}(I)\geqslant p^m_{\theta^m,t}(I)\exp\left(-\frac{\theta^m_I/\sqrt{T}}{p_{\theta^m,t}(I)}\right).$$

Using that $e^{-x}\geqslant 1-x$ for all $x\in\mathbb{R}$, leads to, for all $\theta^m_I\in[r,R]$,

$$p^m_{\theta^m,t+1}(I)\geqslant p^m_{\theta^m,t}(I)-\frac{\theta^m_I}{\sqrt{T}}\geqslant p^m_{\theta^m,t}(I)-\frac{R}{\sqrt{T}}.$$

Summing for $s\in[t]$ leads to,

$$p^m_{\theta^m,t}(I)\geqslant\frac{1}{D_m}-\frac{R}{\sqrt{T}}t\geqslant\frac{1}{D}-\frac{R}{\sqrt{T}}t.$$

Therefore, letting $T_\varepsilon=\lfloor\left(\frac{1}{D}-\varepsilon\right)\frac{\sqrt{T}}{R}\rfloor$ leads to, for all $t\in[T_\varepsilon]$,

$$p^m_{\theta^m,t}(I)\geqslant\frac{1}{D}-\frac{R}{\sqrt{T}}T_\varepsilon\geqslant\varepsilon.$$

F.2.3. *Proof of Proposition 10.* Let $m\in\mathcal{M}$, and let $\theta,\delta\in[r,R]^{D_m}$. The function softmax is 1-Lipschitz with respect to the $\|\cdot\|_2$-norm in $\mathbb{R}^{D_m}$ (see [33] for a proof). Therefore,

$$\|p^m_{\theta,t+1}-p^m_{\delta,t+1}\|_2\leqslant\left\|\sum_{s=1}^t\hat{L}^m_{\theta,s}-\sum_{s=1}^t\hat{L}^m_{\delta,s}\right\|_2\leqslant\sum_{s=1}^t\|\hat{L}^m_{\theta,s}-\hat{L}^m_{\delta,s}\|_2.$$

Thus,

$$\|p^m_{\theta,t+1}-p^m_{\delta,t+1}\|_2\leqslant\frac{1}{\sqrt{T}}\sum_{s=1}^t\left\|\left(\frac{\theta}{p^m_{\theta,s}}-\frac{\delta}{p^m_{\delta,s}}\right)\mathbf{1}_{X_s\in\cdot}\right\|_2$$

where

$$\left\|\left(\frac{\theta}{p^m_{\theta,s}}-\frac{\delta}{p^m_{\delta,s}}\right)\mathbf{1}_{X_s\in\cdot}\right\|_2^2=\sum_{J\in\mathcal{I}^m}\left(\frac{\theta_J}{p^m_{\theta,s}(J)}-\frac{\delta_J}{p^m_{\delta,s}(J)}\right)^2\mathbf{1}_{X_s\in J}.$$

With the triangle inequality, for all $t \in [T_\varepsilon]$

$$
\begin{aligned}
\|p^m_{\theta,t+1} - p^m_{\delta,t+1}\|_2 &\leqslant \frac{1}{\sqrt{T}} \sum_{s=1}^{t} \left\| \left(\frac{\theta}{p^m_{\theta,s}} - \frac{\theta}{p^m_{\delta,s}} \right) \mathbf{1}_{X_s \in \cdot} \right\|_2 + \frac{1}{\sqrt{T}} \sum_{s=1}^{t} \left\| \left(\frac{\theta}{p^m_{\delta,s}} - \frac{\delta}{p^m_{\delta,s}} \right) \mathbf{1}_{X_s \in \cdot} \right\|_2 \\
&\leqslant \frac{R}{\varepsilon^2 \sqrt{T}} \sum_{s=1}^{t} \left\| \left(p^m_{\theta,s} - p^m_{\delta,s} \right) \mathbf{1}_{X_s \in \cdot} \right\|_2 + \frac{1}{\sqrt{T}} \sum_{s=1}^{t} \left\| \left(\frac{\theta}{p^m_{\delta,s}} - \frac{\delta}{p^m_{\delta,s}} \right) \mathbf{1}_{X_s \in \cdot} \right\|_2 \\
&\leqslant \frac{R}{\varepsilon^2 \sqrt{T}} \sum_{s=1}^{t} \left\| p^m_{\theta,s} - p^m_{\delta,s} \right\|_2 + \frac{Rt}{\varepsilon \sqrt{T}} \, \|\theta - \delta\|_\infty \\
&\leqslant \frac{1}{D\varepsilon} \|\theta - \delta\|_\infty + \frac{R}{\varepsilon^2 \sqrt{T}} \sum_{s=1}^{t} \left\| p^m_{\theta,s} - p^m_{\delta,s} \right\|_2 .
\end{aligned}
$$

where

- the first inequality holds because of the triangle inequality,
- the second inequality holds because $x \to 1/x$ is $1/\varepsilon^2$-Lipschitz on $(\varepsilon, 1]$ and for all $J \in \mathcal{I}^m$, $\theta_J \in [r, R]$,
- the third inequality holds because for all $J \in \mathcal{I}^m$, for all $t \in [T_\varepsilon]$ and all $k \in [K]$, $p_{\delta^m,t}(k) \geqslant \varepsilon$ and because $\|\mathbf{1}_{X_s \in \cdot}\|_2^2 = \sum_{J \in \mathcal{I}^m} \mathbf{1}_{X_s \in J} = 1$,
- the last inequality holds since $t \leqslant T_\varepsilon \leqslant \frac{\sqrt{T}}{DR}$.

By the discrete version of Gronwall's Lemma [21], for all $t \in [T_\varepsilon]$,

$$
\|p^m_{\theta,t} - p^m_{\delta,t}\|_2 \leqslant \frac{1}{D\varepsilon} \|\theta - \delta\|_\infty \exp\left(\frac{Rt}{\varepsilon^2 \sqrt{T}} \right) \leqslant \frac{1}{D\varepsilon} \exp\left(\frac{1}{D\varepsilon^2} \right) \|\theta - \delta\|_\infty .
$$

To conclude, note that $\log$ is $1/\varepsilon$-Lipschitz on $[\varepsilon, 1]$ and that $\|p^m_{\theta,t} - p^m_{\delta,t}\|_\infty \leqslant \|p^m_{\theta,t} - p^m_{\delta,t}\|_2$.