Online non-parametric likelihood-ratio estimation by Pearson-divergence functional minimization

Alejandro de la Concha **Nicolas Vayatis** **Argyris Kalogeratos**
Centre Borelli, ENS Paris-Saclay, Université Paris-Saclay, Gif-sur-Yvette, France

Abstract

Quantifying the difference between two probability density functions, p and q, using available data, is a fundamental problem in Statistics and Machine Learning. A usual approach for addressing this problem is the likelihood-ratio estimation (LRE) between p and q, which -to our best knowledge- has been investigated mainly for the offline case. This paper contributes by introducing a new framework for online non-parametric LRE (OLRE) for the setting where pairs of iid observations $(x_t \sim p, x'_t \sim q)$ are observed over time. The non-parametric nature of our approach has the advantage of being agnostic to the forms of p and q. Moreover, we capitalize on the recent advances in Kernel Methods and functional minimization to develop an estimator that can be efficiently updated online. We provide theoretical guarantees for the performance of the OLRE method along with empirical validation in synthetic experiments.

1 INTRODUCTION

The likelihood-ratio between two probability density functions (pdfs) is a quantity omnipresent in Statistics. For instance, the likelihood-ratio test has optimal statistical power and it is a core tool in statistical hypothesis testing (Neyman and Pearson, 1933; Casella and Berger, 2006). In one of the related problems, change-point detection, the most widely-used methods, such as CUSUM (Page, 1954) or Sriryaev-Roberts (Shiryaev, 1963), depend on the likelihood-ratio (see also Tartakovsky et al. (2014); Xie et al. (2021)). In Transfer Learning, it is possible to define a weighted cost function to solve a new problem taking into account prior knowledge provided by a different dataset; interestingly, this weighting function coincides with the likelihood-ratio (Fishman, 1996; Zhuang et al., 2021).

ϕ-divergences are also ubiquitous in Statistics. Classical problems such as Maximum Likelihood Estimation, Dimensionality Reduction, and Generative Modeling, to mention just a few, can be restated as ϕ-divergences minimization problems (Liese and Vajda, 2006; Nguyen et al., 2010; Sugiyama et al., 2012; Agrawal and Horel, 2021).

From a Machine Learning perspective, the interplay between the likelihood-ratio and ϕ-divergences has been described via its variational representation (Nguyen et al., 2008). There were identified situations where the ϕ-divergence estimation between two measures amounts to a likelihood-ratio estimation (LRE) *as an element* in a functional space. This kind of result has motivated a plethora of non-parametric methods, based on Kernel Methods and Neural Networks (Moustakides and Basioti, 2019), which do not need any further hypotheses regarding the functional form of p and q and just depend on observations coming from both those probability densities. These techniques have a wide range of applications in different domains (Basseville, 2013; Liu et al., 2013; Rubenstein et al., 2019; Zhang and Yang, 2021).

Despite the aforementioned success of non-parametric LRE methods for offline processing, to our best knowledge, there has been hardly any investigation about how the estimation of ϕ-divergence and likelihood-ratio can be adapted to online settings and streaming data. The motivation for covering this gap is to pave the way so that non-parametric LRE-based methods bring gains in online learning, hypothesis testing, or various detection settings.

Contribution. To begin with, in this paper we introduce the new Online LRE (OLRE) problem, where one observes a stream of incoming pairs of observations $(x_t \sim p,\ x'_t \sim q)$, $t = 1, 2, ...$, and the likelihood-ratio

Proceedings of the 27[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238.

needs to be estimated on the fly. Then, we present the homonymous non-parametric OLRE framework, along with a theoretical characterization of its convergence. Our approach nurtures mainly from three elements:

- The formulation of the typical offline LRE problem as a functional minimization problem seeking a solution among functions of a Reproducing Kernel Hilbert Space (RKHS) (Nguyen et al., 2008).
- The adaptation of a first-order optimization method, the stochastic functional gradient descent (Kivinen et al. (2004)), combined with the framework of regularized paths in Hilbert spaces (Tarrès and Yao, 2014).
- The organic integration of up-to-date practices for kernel-based LRE.

The OLRE framework combines the above elements, and enjoys the following technical properties:

- It does not require to know in advance the sample size, which can be even infinite.
- Our stochastic approximation aims to minimize the generalization error by solving the original functional minimization problem, instead of performing empirical risk minimization, hence avoids over-fitting.
- Our analysis and the performance of the proposed method, highlight the bias of existing offline approaches that are based on empirical risk minimization, which rely on simple heuristics to manage large amounts of data.
- The cost of the iteration at time t is $O(t)$, hence in total $O(t^2)$ for up to time t.
- Our convergence results provide guidelines on how to select the hyperparameters of our method and its sensibility to different configurations.

2 PROBLEM STATEMENT AND BACKGROUND

In this section, we begin by presenting the likelihood-ratio estimation (LRE) problem, and by defining the Online LRE (OLRE) problem version. Then, we present the main building blocks we use for developing the homonymous OLRE framework.

2.1 Likelihood-ratio estimation

Let us denote the feature space $\mathcal{X} \subset \mathbb{R}^d$ and consider two probability measures P and Q which are absolutely continuous with respect to the Lebesgue measure denoted by dx, with densities p and q respectively. We also define the convex α-mixture of the probability measures P and Q computed by $P^\alpha = (1-\alpha)P + \alpha Q$, and similarly for their densities p, q, where $0 \leq \alpha < 1$ is user-defined parameter.

Relative likelihood-ratio. We focus on the approximation of the relative likelihood-ratio between the pdfs q and p:

$$r^\alpha(x) = \frac{q(x)}{(1-\alpha)p(x) + \alpha q(x)} \in \mathbb{R}^+, \quad \forall x \in \mathcal{X}, \quad (1)$$

where $0 \leq \alpha < 1$ acts as a user-defined regularization parameter (Yamada et al., 2011). When $\alpha = 0$, Eq. 1 recovers the usual likelihood-ratio $r_*^{\alpha=0}(x) = r(x) = \frac{q(x)}{p(x)}$. The α-regularization addresses certain instability issues appearing when approximating an unbounded function. Specifically, when $\alpha > 0$, it holds $r^\alpha \leq \frac{1}{\alpha}$, which is an upper-bound that will be proven to be important when we later study theoretically the convergence of the proposed method (see Sec. 4). Typically, α should be close to 0 to ensure that the approximated $r^\alpha(x)$ will remain relevant for the intended application of the likelihood-ratio, which is of course the core quantity of interest.

Defining the Online LRE (OLRE) setting. The online setting we introduce in this paper supposes that a new pair of iid observations ($x_t \sim p$, $x'_t \sim q$) is observed at every time t. Then, the objective is to approximate the relative likelihood-ratio r^α through a function f_t, which is updated at every time t. The function f_t is an element of a non-parametric functional space, so there is no need to make a hypothesis about the nature of p nor q. We denote by Ξ_t the minimum σ-algebra generated by the incoming observations up to time t, i.e. $\sigma(\{(x_1, x'_1), ..., (x_t, x'_t)\})$.

Reproducing Kernel Hilbert Space. We aim to estimate $r^\alpha(x)$ with regards to a Reproducing Kernel Hilbert Space (RKHS) $\mathbb{H}$ containing as elements functions $f : \mathcal{X} \to \mathbb{R}$. $\mathbb{H}$ is equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}} : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$, which will be reproduced by a Mercer Kernel; i.e. by a continuous symmetric real function, which is the positive semi-definite kernel function $K(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. Then, the space $\mathbb{H}$ satisfies the following properties:

- $\langle K(x, \cdot), f \rangle_{\mathbb{H}} = f(x)$, for any $f \in \mathbb{H}$;
- $\mathbb{H} = \overline{\text{span}}(\{K(x, \cdot) : \forall x \in \mathcal{X}\})$, (2)

where $\overline{\text{span}}$ refers to the closure of all the linear combinations of the elements $K(x, \cdot)$, $\forall x \in \mathcal{X}$. The first equality is known as the RKHS reproducing property.

2.2 Important notions for first-order optimization

The main idea behind OLRE is to see r^α as the solution of the functional optimization problem $\min_{f \in \mathbb{H}} L(f)$, where $L(f) : \mathbb{H} \to \mathbb{R}$ is a cost functional representing the real risk (i.e. generalization error), with respect to an instantaneous loss-function

$\ell(f): \mathbb{H} \to \mathbb{R}$. Our optimization schema is based on the functional gradient of the cost function $\ell(f)$, and produces a stochastic approximation f_t that approaches the relative likelihood-ratio r^α at every time t. The estimation of f_t requires only the previous estimate f_{t-1} and the new observations (x_t, x'_t). The geometry of $\mathbb{H}$, and more precisely the reproducing property of its elements, lead to an elegant closed-form expression for f_t.

***Functional gradient* (Bauschke and Combettes, 2011).** Let $L: \mathbb{H} \to \mathbb{R}$ be a Gâteaux differentiable functional, and $[DL(f)](\cdot)$ its Gâteaux derivative. By $\nabla_f L(f)$ we denote the functional gradient of L at f, defined to be the element of $\mathbb{H}$ that satisfies:

$$[DL(f)](g) = \langle \nabla_f L(f), g \rangle_{\mathbb{H}}, \quad \forall g \in \mathbb{H}. \tag{3}$$

The Riesz representation theorem tells us that $\nabla_f L(f)$ exists and is unique. When $L(f)$ is also Fréchet differentiable at f, then the Gâteaux derivative and Fréchet derivative coincide, but the latter has the advantage that satisfies the chain rule in a more natural way.

***Functional Stochastic Gradient* (Kivinen et al., 2004).** Suppose that the cost function $L(f)$ takes values in an RKHS, i.e. $f \in \mathbb{H}$, and it has the form $L(f) = \mathbb{E}_x[\ell(f(x)]$. Given an independent realization $x \in \mathcal{X}$, we can compute the Fréchet derivative of $\ell(f(\cdot))$ w.r.t. f as:

$$\nabla_f \ell(f(x))(\cdot) = \frac{\partial \ell(f(x))}{\partial f(x)} \frac{\partial f(x)}{\partial f}(\cdot) = \frac{\partial \ell(f(x))}{\partial f(x)} K(x, \cdot). \tag{4}$$

The first equality is a consequence of the chain rule for the Fréchet derivative; the second one is due to $\mathbb{H}$'s reproducing property that due to which $\frac{\partial f(x)}{\partial f}(\cdot) = \frac{\partial \langle f, K(x,\cdot)\rangle_{\mathbb{H}}}{\partial f}(\cdot) = K(x, \cdot)$. The operator $\nabla_f \ell(f(x))(\cdot)$ is the functional stochastic gradient of $L(f)$ at f.

3 ONLINE LRE

3.1 LRE via ϕ-divergence minimization

***ϕ-divergence*.** Under the assumptions mentioned at the beginning of Sec. 2.1 regarding P and Q, a ϕ-divergence functional quantifies the dissimilarity between two probability measures that are described by their pdfs p, q:

$$\mathcal{D}_\phi(P\|Q) = \int \phi\left(\frac{q(x)}{p(x)}\right) p(x)\, dx = \int \phi(r^*)(x)\, dP(x), \tag{5}$$

where $\phi: \mathbb{R} \to \mathbb{R}$ is a convex and semi-continuous real function with $\phi(1) = 0$ (Csiszár, 1967). This definition assumes absolute continuity between Q and P, that is $P \ll Q$.

The formulation of our optimization problem relies mainly on the following variational representation for ϕ-divergences.

Lemma 1. *(Lemma 1 in Nguyen et al. (2008)). Consider P and Q two probability measures satisfying the assumptions listed in Sec. 2.1. Then, if P is absolutely continuous with respect to Q, for any class of functions $\mathcal{F}: \mathcal{X} \to \mathbb{R}$, the lower-bound for the similarity between two probability measures is:*

$$\mathcal{D}_\phi(P\|Q) = \int \phi\left(\frac{q}{p}\right)(x)\, dP(x) \tag{6a}$$

$$\geq \sup_{g \in \mathcal{F}} \int g(x') dQ(x') - \int \phi^\star(g)(x)\, dP(x) \tag{6b}$$

where $\phi^\star$ denotes the convex conjugate of $\phi: \mathbb{R} \to \mathbb{R}$. The equality Eq. 6a holds if and only if the subdifferential $\nabla\phi(\frac{q}{p})$ contains an element of $\mathcal{F}$.

The likelihood-ratio in terms of the solution to Problem 6b, $g^* = \sup_{g \in \mathcal{F}} \int g(x')dQ(x') - \int \phi^\star(g)(x)dP(x)$, can be inferred by simply applying $\frac{q}{p} = (\nabla\phi)^{-1}(g^*) = \nabla\phi^\star(g^*)$. For this to be possible, ϕ needs to be continuously-differentiable and strictly convex. As stated in the introduction, though, we will focus on approximating the relative likelihood-ratio r^α instead of the usual likelihood-ratio. Then, by fixing $\phi(y) = \frac{(y-1)^2}{2}$ for $y \in \mathbb{R}$, whose convex conjugate is $\phi^\star(y^*) = \frac{(y^*)^2}{2} + y^*$ for $y^* \in \mathbb{R}$, we recover the χ^2-divergence (also known as Pearson-divergence):

$$PE(P^\alpha\|Q) = \int \left[\frac{(r^\alpha - 1)^2}{2}\right](x)\, dP^\alpha(x). \tag{7}$$

Note that the factor $\frac{1}{2}$ is only introduced to facilitate later calculations. According to Lemma 1, the latter can be lower-bounded via its variational representation:

$$\begin{aligned} &PE(P^\alpha\|Q) \\ &\quad \geq \sup_{f \in \mathbb{H}} \int (f-1)(x')\, dQ(x') \\ &\qquad - \int \left[\frac{(f-1)^2}{2} + (f-1)\right](x)\, dP^\alpha(x) && \text{(8a)} \\ &\quad = \sup_{f \in \mathbb{H}} \int f(x')\, dQ(x') - \int \frac{f^2(x)}{2}\, dP^\alpha(x) - \frac{1}{2} && \text{(8b)} \\ &\quad = \sup_{f \in \mathbb{H}} \int \left[f r^\alpha - \frac{f^2}{2}\right](x)\, dP^\alpha(x) - \frac{1}{2}. && \text{(8c)} \end{aligned}$$

Eq. 8b is explained by the change of measure identity $\mathbb{E}_{p^\alpha(y)}[f(y) r^\alpha(y)] = \mathbb{E}_{q(x')}[g(x')]$. Then, thanks to Lemma 1, we can obtain an estimator of the r^α by solving the following quadratic functional optimization problem defined in terms of the RKHS $\mathbb{H}$:

$$
\begin{aligned}
&\underset{f\in\mathbb{H}}{\operatorname{argmin}}\, L^{\mathrm{PE}}(f) \\
&= \underset{f\in\mathbb{H}}{\operatorname{argmin}} \int \left[\frac{f^2(x)}{2} - f(x)r^\alpha(x)\right] dP^\alpha(x) + \frac{1}{2} \quad (9a)\\
&= \underset{f\in\mathbb{H}}{\operatorname{argmin}} \int \frac{(f-r^\alpha)^2(x)}{2} dP^\alpha(x) + C \quad (9b)\\
&= \underset{f\in\mathbb{H}}{\operatorname{argmin}}\ (1-\alpha)\int \frac{f^2(x)}{2} dP(x) + \alpha \int \frac{f^2(x')}{2} dQ(x') \\
&\qquad - \int f(x')\, dQ(x') + C. \quad (9c)
\end{aligned}
$$

To get Eq. 9b, we used that $\mathbb{E}_{p^\alpha(y)}[r^\alpha(y)^2] = C$, where C is some constant. For the final Eq. 9c, we used $\mathbb{E}_{p^\alpha(y)}[f(y)] = \alpha \mathbb{E}_{p(x)}[f(x)] + (1-\alpha)\mathbb{E}_{q(x')}[f(x')]$.

In the rest of this section, we present our online LRE method. Later, in Sec. 5, we provide a discussion that compares our approach with existing works that aim to solve the LRE problem in an offline fashion.

3.2 Online LRE by χ^2-divergence minimization

In the previous section, we put forward our goal to solve the functional optimization Problem 9 as new pairs of iid observations ($x_t \sim p, x'_t \sim q$) arrive over time. The proposed homonymous algorithm, which is called as well OLRE, makes use of the approach of regularization paths.

Let us define next the regularized cost function with the help of a time-dependent regularization parameter $\lambda_t > 0$:

$$
\begin{aligned}
\min_{f\in\mathbb{H}}\ &(1-\alpha)\int \frac{f^2(x)}{2} dP(x) + \alpha \int \frac{f^2(x')}{2} dQ(x') \\
&- \int f(x')\, dQ(x') + \frac{\lambda_t}{2} \|f\|_{\mathbb{H}}
\end{aligned} \quad (10)
$$

The idea of stochastic approximation via regularization paths is to use a decreasing regularization sequence $\{\lambda_t \in \mathbb{R}^+\}_{t\in\mathbb{N}}$ in the regularized Problem 10 in order to generate a sequence of estimated relative likelihood-ratios $\{f_t\}_{t\in\mathbb{N}}$ that converges to the target: $f_t \to r^\alpha$ as $\lambda_t \to 0$.

To describe precisely the stochastic approximation strategy for the online setting, we first define the regularized instantaneous cost function $\ell_t^{\mathrm{PE}}(f)$, $f \in \mathbb{H}$, based on Eq. 10:

$$
\ell_t^{\mathrm{PE}}(f) = (1-\alpha)\frac{f^2(x_t)}{2} + \alpha\frac{f^2(x'_t)}{2} - f(x'_t) + \frac{\lambda_t}{2}\|f\|_{\mathbb{H}}^2. \quad (11)
$$

The functional stochastic gradient $\nabla_f(\ell_t^{\mathrm{PE}}(f))(\cdot)$ gives the random direction of the stochastic update. Thanks to its properties listed in Sec. 2.2, we can compute it easily by:

$$
\begin{aligned}
&\nabla_f(\ell_t^{\mathrm{PE}}(f))(\cdot) \\
&= \frac{(1-\alpha)}{2}\nabla_f\left(f^2(x_t)\right)(\cdot) + \frac{\alpha}{2}\nabla_f\left(f^2(x'_t)\right)(\cdot) \\
&\quad - \nabla_f(f(x'_t))(\cdot) + \lambda_t f(\cdot) \\
&= (1-\alpha)f(x_t)K(x_t,\cdot) + (\alpha f(x'_t) - 1)K(x'_t,\cdot) + \lambda_t f(\cdot).
\end{aligned} \quad (12)
$$

where the last equality is a consequence of Expr. 4.

Let us denote by $SL(\mathbb{H})$ be the set of self-adjoint bounded linear operators in $\mathbb{H}$. Then, we can define the random variables $A_t : \mathcal{X}\times\mathcal{X} \to SL(\mathbb{H})$ and $b_t : \mathcal{X}\times\mathcal{X} \to \mathbb{H}$ as:

$$
\begin{aligned}
A_t &= (1-\alpha)\langle\cdot, K(x_t,\cdot)\rangle_{\mathbb{H}} K(x_t,\cdot) \\
&\quad + \alpha\langle\cdot, K(x'_t,\cdot)\rangle_{\mathbb{H}} K(x'_t,\cdot) + \lambda_t I_{\mathbb{H}} \\
&= A(x_t, x'_t) + \lambda_t I_{\mathbb{H}} \\
b_t &= K(x'_t,\cdot),
\end{aligned} \quad (13)
$$

where $I_{\mathbb{H}}$ is the identity operator in $\mathbb{H}$, and $A : \mathcal{X}\times\mathcal{X} \to SL(\mathbb{H})$ is such that when applied to $f \in \mathbb{H}$:

$$
A(x,x')f = (1-\alpha)f(x)K(x,\cdot) + \alpha f(x')K(x',\cdot). \quad (14)
$$

Then, the functional stochastic gradient can be rewritten in terms of A_t and b_t as:

$$
\nabla_f(\ell_t^{\mathrm{PE}}(f))(\cdot) = A(x_t,x'_t)f + \lambda_t f(\cdot) - b_t = A_t f - b_t, \quad (15)
$$

and the stochastic update for Problem 9 becomes:

$$
\begin{aligned}
f_t(\cdot) &= f_{t-1}(\cdot) - \eta_t \nabla_f(\ell_t^{\mathrm{PE}}(f_{t-1}))(\cdot) \\
&= (1-\eta_t\lambda_t)f_{t-1}(\cdot) - \eta_t[A(x_t,x_t)f_{t-1} - b_t] \\
&= (1-\eta_t\lambda_t)f_{t-1}(\cdot) - \eta_t[(1-\alpha)f_{t-1}(x_t)K(x_t,\cdot) \\
&\quad + (\alpha f_{t-1}(x'_t) - 1)K(x'_t,\cdot)],
\end{aligned} \quad (16)
$$

where $\eta_t > 0$ is a given step-size at time t. We will discuss in Sec. 4 which are the conditions to be satisfied by the sequence $\{\eta_t\}_{t\in\mathbb{N}}$ and $\{\lambda_t\}_{t\in\mathbb{N}}$ so that f_t converges.

Suppose a dictionary D_{t-1} made of M_{t-1} basis functions, $\{x_1, ..., x_{M_{t-1}} \in \mathbb{H}\}$, and a kernel function $K : \mathcal{X} \to \mathbb{R}^{M_{t-1}}$ that maps input data to vectors: $K(D_{t-1},\cdot) = (K(x_1,\cdot), ..., K(x_{M_{t-1}},\cdot))^\top$. Then, if we express $f_{t-1}(\cdot)$ using D_{t-1} and a weight vector $\theta_{t-1} \in \mathbb{R}^{M_{t-1}}$:

$$
f_{t-1}(\cdot) = \sum_{m=1}^{M_{t-1}} \theta_{t-1,m} K(x_m,\cdot) = K(D_{t-1},\cdot)^\top \theta_{t-1}, \quad (17)
$$

we can also express the subsequent f_t using the extended dictionary $D_t = D_{t-1} \cup \{K(x_t,\cdot), K(x'_t,\cdot)\}$. In that case, the new weights come from the concatenation of the previous weights and two new terms

Algorithm 1: Online LRE (OLRE)

Input : $\{x_t \sim p, x'_t \sim q\}_{t=1,\ldots}$: stream of observation pairs;
t_0: size of the warm-up period;
$a \geq 4$, $\frac{1}{2} \leq \beta \leq 1$: fixed constants
$0 < \alpha < 1$: prefixed regularization parameter;
K: predefined kernel (form and hyperparameters).
Output: $\{f_t\}_{t=1}^{T}$: set of estimated relative likelihood-ratios.

Initialize $f_0(\cdot) = 0$; $D_0 = []$; $\theta_0 = []$
for $t = 1, 2, \ldots$ **do**
  Get the incoming iid pair of observations (x_t, x'_t)
  Compute the step-size and the penalization parameter:

$$\eta_t = a\left(\frac{1}{t_0 + t}\right)^{\frac{2\beta}{2\beta+1}}, \quad \lambda_t = \frac{1}{a}\left(\frac{1}{t_0 + t}\right)^{\frac{1}{2\beta+1}} \tag{18}$$

  Update the dictionary:
    $D_t = D_{t-1} \cup \{x_t, x'_t\}$
  Update the weights:
    $\theta_t = [(1-\eta_t\lambda_t)\theta_{t-1}, \eta_t(\alpha-1)f_{t-1}(x_t), \eta_t(1-\alpha f_{t-1}(x'_t))]$
  Update the relative likelihood-ratio estimate:
    $f_t(\cdot) = K(D_t, \cdot)^{\top}\theta_t$
end
return $\{f_t\}_{t=1}^{T}$

depending on f_{t-1} evaluated at x_t and x'_t: $\theta_t = [(1-\eta_t\lambda_t)\theta_{t-1}, \eta_t(\alpha-1)f_{t-1}(x_t), \eta_t(1-\alpha f_{t-1}(x'_t))] \in \mathbb{R}^{M_{t-1}+2}$.

The relationship between f_{t-1} and f_t implies that the cost per iteration is mainly for computing $f_{t-1}(x_t)$, which requires $2(t-1)$ kernel function evaluations. Therefore, the cost per iteration scales rate $\mathcal{O}(t)$, and that the number of kernel function evaluations up to time t is $\mathcal{O}(t^2)$. A sketch of the OLRE algorithm is shown in Alg. 1. The implementation of OLRE and the experimental scenarios tested in this paper are publicly available online[1].

4 THEORETICAL GUARANTEES

Previous convergence analyses of LRE are restricted to the offline setting where n pairs of observations from p and n' q are available at the time of estimation. Works such as Sugiyama et al. (2007); Nguyen et al. (2008, 2010); Yamada et al. (2011); Sugiyama et al. (2012), capitalized over available theoretical results for M-estimators (van de Geer, 2000). That framework is successfully adapted to derive convergence rates as most of the LRE rely on a penalized cost function based on an empirical approximation of ϕ-divergences. The metrics that were used to describe the convergence of the likelihood-ratio estimates, which we will denote by $\hat{f}_{\lambda_n}$, depend on the ϕ-divergence that is used for estimation. More precisely, it is common to define an estimator $D_n(\hat{f}_{\lambda_n})$ aiming to approximate the real ϕ-divergence $\mathcal{D}_\phi(P\|Q)$ (Eq. 5) to then describe the convergence of the method via an upper-bound of the quantity $\left|D_n(\hat{f}_{\lambda_n}) - \mathcal{D}_\phi(P\|Q)\right|$. It is common as well to derive convergence rates in terms of a similarity measure between $\hat{f}_{\lambda_n}$ and the real likelihood-ratio r; the similarity measure is chosen as well based on the ϕ-divergence. For example, Sugiyama et al. (2007) and Nguyen et al. (2008, 2010) study the LRE problem based on the Kullback-Leibler divergence, and the convergence rates between $\hat{f}_{\lambda_n}$ and r are given in terms of the Hellinger distance.

The M-estimation approach requires further hypotheses over the functional space $\mathcal{F}$ and the real likelihood-ratio function r. For example, the convergence rates depend on the complexity of $\mathcal{F}$ summarized in quantities such as covering numbers or bracketing numbers. It is common to set unrealistic assumptions over the real likelihood-ratios, such as a strictly positive lower-bound and a finite upper-bound even when r is unregularized (Nguyen et al., 2008, 2010; Sugiyama et al., 2012). Moreover, although those existing approaches and theoretical results assume that all observations are used in the estimation process, their associated numerical implementations require fixing a finite-dimensional dictionary, whose impact on those convergence rates has not been detailed.

Theorems 1 and 2 summarize the OLRE convergence rates in terms of the $L^2_{p^\alpha}$ and the Hilbert norms. The theoretical approach used to produce these results differs from previous works as we deal directly with the functional optimization problem described in Lemma 1 without using the empirical risk as a surrogate cost function, nor the hypothesis of a fixed number of observations (i.e. fixed horizon). This implies that the proofs of both theorems (see Appendix B) no longer depend on M-estimation nor the required restrictive hypotheses of that approach. Instead, we employ stochastic approximation of regularized paths (Tarrès and Yao, 2014), which deals with the online solution of a linear operator equation defined in a Hilbert space $\mathbb{H}$. In fact, we show in the appendix how the Pearson-based optimization of Problem 9 is connected with the regression problem in $\mathbb{H}$ as both can be written as linear operator equations in an RKHS. This stochastic approach allows us to obtain for the first time convergence rates in terms of the Hilbert norm and with milder hypotheses. Furthermore, as we use all the observations in the numerical implementation, there is no gap between theory and practice regarding the convergence rates analyzed in both theorems.

4.1 Convergence guarantees for OLRE

Covariance operator. The covariance operator is a key component for studying OLRE's convergence properties (see Appendix B). Let $\mathcal{L}^2_{p^\alpha}$ be the space of square integrable functions with respect to p^α, and $L^2_{p^\alpha}$ its quotient space, which is a Hilbert space whose

[1] https://github.com/AlejandrodelaConcha/OLRE

norm is denoted by $\|\cdot\|_{L^2_{p^\alpha}}$. Notice that if p^α has full support on $\mathcal{X}$, then we can do the usual identification of the elements of $\mathcal{L}^2_{p^\alpha}$ and its equivalent classes in $L^2_{p^\alpha}$.

Let us denote by $\mathcal{L}_K : L^2_{p^\alpha} \to L^2_{p^\alpha}$ the linear operator defined by the following integral transform:

$$\mathcal{L}_K(f)(t) = \int_{\mathcal{X}} K(t,x) f(x)\, dP^\alpha(x). \tag{19}$$

The operator $\mathcal{L}_K$ has been studied in detail in Dieuleveut and Bach (2016). $\mathcal{L}_K$ is a bounded self-adjoint semi-definite positive operator on $L^2_{p^\alpha}$ and it is trace-class. Furthermore, it is possible to show that there exists an orthonormal eigensystem $\{\mu_k, \psi_k\}_{k\in\mathbb{N}}$ in $L^2_{p^\alpha}$, where μ_k is a basis of $\mathbb{H}$, and that the eigenvalues $\{\mu_k\}_{k\in\mathbb{N}}$ are strictly positive and arranged in decreasing order (see Proposition 2.2 in Dieuleveut (2017)). The eigen-elements can be used to define the operator $\mathcal{L}^\beta_K : L^2_{p^\alpha} \to L^2_{p^\alpha}$, for $\beta \in \mathbb{R}$:

$$\mathcal{L}^\beta_K\left(\sum_{k\in\mathbb{N}} c_k \psi_k\right) = \sum_{k\in\mathbb{N}} c_k \mu_k^\beta \psi_k. \tag{20}$$

The operator $\mathcal{L}^\beta_K$ is relevant as it encodes how well the chosen kernel approximates the relative likelihood-ratio. More precisely, the norm $\|\mathcal{L}^\beta_K r^\alpha\|_{\mathbb{H}}$ defines a notion of smoothness of r^α w.r.t. $\mathbb{H}$. In particular, for $\beta = \frac{1}{2}$, $\mathcal{L}^{\frac{1}{2}}_K$ defines an isometric isomorphism of Hilbert spaces (see Proposition 3 in Dieuleveut and Bach (2016)), that is $\|f\|_{L^2_{p^\alpha}} = \left\|\mathcal{L}^{\frac{1}{2}}_K f\right\|_{\mathbb{H}}$.

When $\mathcal{L}_K$ is restricted to elements $f \in \mathbb{H} \subset L^2_{p^\alpha}$, we recover the covariance operator, which is known to satisfy that $\forall f, g \in \mathbb{H}$, $\langle f, \mathcal{L}_K(g)\rangle_{\mathbb{H}} = \mathbb{E}_{p^\alpha(y)}[f(y)g(y)]$.

Main convergence results

Assumption 1. *The pairs of observations $(x_t, x'_t), t = 1, 2, \dots$ are iid in time and satisfy $x_t \sim p$ and $x'_t \sim q$.*

The independence hypothesis is present in Nguyen et al. (2008, 2010) and in the general theoretical framework for LRE of Sugiyama et al. (2012).

Assumption 2. *The reproducing kernel map can be upper-bounded by a constant $C > 0$: $\sup_{x\in\mathcal{X}} \sqrt{K(x,x)} \le C < \infty$.*

This assumption allows to bound the functions $f \in \mathbb{H}$ in terms of the $\|\cdot\|_{\mathbb{H}}$. It is satisfied by commonly used kernels, such as the Gaussian and the Laplacian kernels, and in general for any continuous $K(\cdot,\cdot)$ defined in a compact input feature space $\mathcal{X}$.

Assumption 3. *p^α has full support on the feature space $\mathcal{X}$.*

This statement enhances the use of the covariance operator (Dieuleveut and Bach, 2016) and it is an important hypothesis for the framework presented in Tarrès and Yao (2014).

Assumption 4. *$r^\alpha \in L^\beta_K(\mathcal{L}^2_{p^\alpha})$ for $\frac{1}{2} \le \beta \le 1$.*

The parameter β controls the smoothness of r^α in $\mathbb{H}$. Assumption 4 implies that the proposed model is well-defined, in the sense that $r^\alpha \in \mathbb{H}$, which is the usual hypothesis made in the LRE literature (Sugiyama et al., 2012). Moreover, as β increases, $\mathcal{L}^\beta_K(L^2_{p^\alpha})$ defines a sequence of decreasing subspaces of $L^2_{p^\alpha}$, i.e. higher β values assume a stronger smoothness of r^α.

Theorem 1 gives OLRE's convergence with respect to the space $\mathcal{L}^2_{p^\alpha}$. The norm $\|f_t - r^\alpha\|^2_{\mathcal{L}^2_{p^\alpha}}$ equals to the real least-squared error $\mathbb{E}_{p^\alpha(y)}[(f_t - r^\alpha)^2(y)]$. Moreover, this convergence result can be easily applied to describe the convergence with respect to the excess risk $L^{\text{PE}}(f) - L^{\text{PE}}(r^\alpha)$.

Theorem 1. *(**OLRE's convergence in $\mathcal{L}^2_{p^\alpha}$**) Given Assumptions 1-4, $a \ge 4$ and $t_0 \ge (2 + 4C^2a)^{\frac{(2\beta+1)}{2\beta}}$. Then if the learning rate sequence is fixed as $\eta_t = a\left(\frac{1}{t}\right)^{\frac{2\beta}{2\beta+1}}$ and $\lambda_t = \frac{1}{a}\left(\frac{1}{t}\right)^{\frac{1}{2\beta+1}}$. Then for all $t \in \mathbb{N}$ and $\delta \in (0,1)$, with probability at least $1-\delta$:*

$$\|f_t - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} \le \frac{C_1}{\bar{t}} + \left(C_2 a^{(-\beta)} + C_3\sqrt{a}\log(\frac{2}{\delta})\right)\left(\frac{1}{\bar{t}}\right)^{\frac{\beta}{2\beta+1}} + \left(C_4 a^{\frac{5}{2}} + C_5 a^{\frac{7}{2}}\sqrt{log(\bar{t})}\right)(\log^2(\frac{2}{\delta}))\left(\frac{1}{\bar{t}}\right)^{\frac{4\beta-1}{4\beta+2}}, \tag{21}$$

where:

$$C_1 = \frac{2t_0}{\alpha}, C_2 = \frac{5\beta+1}{\beta(1+\beta)}\left\|L_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}, C_3 = \frac{16C}{\alpha}$$
$$C_4 = \frac{32C^3}{\alpha}, \quad C_5 = \frac{8C^3(10C+3)}{\alpha}.$$

Notice that the convergence rate in $\mathcal{L}^2_{p^\alpha}$ can be decomposed into three terms. The first depends on the initialization, and decreases at rate $\mathcal{O}(t^{-1})$. The second one is related to the smoothness of the likelihood-ratio in $\mathbb{H}$ and the noise in the observations, and decreases at rate $\mathcal{O}(t^{-\frac{2}{2\beta+1}})$. The third term is related to the variance of the observations, and decreases at a rate $\mathcal{O}(\log^{\frac{1}{2}}(t) t^{-\frac{4\beta-1}{4\beta+2}})$. When $\beta \in (\frac{1}{2}, 1]$, the second term becomes dominant, which implies a faster convergence as the smoothness of r^α increases. When $\beta = \frac{1}{2}$, the convergence rate becomes $\mathcal{O}(\log^{\frac{1}{2}}(t) t^{-\frac{1}{4}})$.

The convergence rate with respect to $\mathbb{H}$ is more restrictive than in $\mathcal{L}^2_{p^\alpha}$. For $\mathbb{H}$, Assumption 4 needs to be replaced by Assumption 5; the main difference is that r^α is required to be smoother with respect to $\mathbb{H}$ for higher β values.

Assumption 5. *$r^\alpha \in L^\beta_K(\mathcal{L}^2_{p^\alpha})$ for $\frac{1}{2} < \beta \le \frac{3}{2}$.*

Theorem 2. *(**OLRE's convergence in** $\mathbb{H}$) Given Assumptions 1-3 and 5, $a \ge 4$ and $t_0 \ge (aC^2+1)^{\frac{(2\beta+1)}{2\beta}}$. Then if the learning rate sequence is fixed as $\eta_t =$*

$a\left(\frac{1}{t+t_0}\right)^{\frac{2\beta}{2\beta+1}}$ *and* $\lambda_t = \frac{1}{a}\left(\frac{1}{t+t_0}\right)^{\frac{1}{2\beta+1}}$. *Then for all* $t \in \mathbb{N}$ *and* $\delta \in (0,1)$, *with probability at least* $1-\delta$:

$$\|f_t - r^\alpha\|_{\mathbb{H}} \leq \frac{C_1'}{\bar{t}} + \left(C_2' a^{\frac{1}{2}-\beta} + C_3' a \log\left(\frac{2}{\delta}\right)\right)\left(\frac{1}{\bar{t}}\right)^{\frac{2\beta-1}{4\beta+2}}, \tag{22}$$

where $\bar{t} = t + t_0$ *and,*

$$C_1' = \frac{2\sqrt{a} t_0^{\frac{4\beta+1}{4\beta+2}}}{\alpha}, C_2' = \frac{20\beta-2}{(2\beta-1)(2\beta+3)}\left\|L_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}$$

$$C_3' = 6\left(\frac{(C+1)^2}{C\alpha}\right).$$

We can see that the upper-bound appearing in Theorem 2 is made of two components. The first component is related to the constant C_1' and summarizes the impact of the initialization. This term converges at rate $\mathcal{O}(t^{-1})$). The second term, which is the leading term of the expression, converges at rate $\mathcal{O}(t^{-\frac{2\beta-1}{4\beta+2}})$ and it mainly depends on the smoothness parameter β. The bigger β , the faster the convergence. Notice that the case $\beta = \frac{1}{2}$ is not considered in this theorem, in fact, the algorithm may not converge in $\mathbb{H}$, as indicated by Theorem A in Tarrès and Yao (2014).

Both Theorems 2 and 1 provide useful information on how to fix the step sizes $\{\eta_t\}_{t\in\mathbb{N}}$ and regularization constants $\{\lambda_t\}_{t\in\mathbb{N}}$, and explain their its impact to the convergence rates. Notice that there is an interplay between the selection of a and the smoothness parameter β. The results suggest that OLRE converges faster in $\mathcal{L}^2_{p^\alpha}$ than in $\mathbb{H}$ if the hyperparameters are the same. Both results shed light on the impact of the parameter α, as values close to 1 will lead to better convergence rates. Nevertheless, $\alpha = 1$ render r^α a constant, which is meaningless for most applications. For this reason, the value of α should take into account both the convergence rate of the optimization schema and the intended application.

Convergence results for likelihood-ratio estimates based on the Pearson-divergence can be found in Yamada et al. (2011). Those results are given in terms of the difference between the real Pearson-divergence $L^{\text{PE}}(r^\alpha)$ and an empirical approximation $L_n^{\text{PE}}(\hat{f}_{\lambda_n})$. It was shown that if the regularization constant decreases at speed $\lambda_n = \mathcal{O}(n^{-\frac{2}{2+\gamma}})$, where the parameter $\gamma \in (0,2)$ quantifies the complexity of $\mathbb{H}$, then RULSIF could achieve a convergence rate $L^{\text{PE}}(r^\alpha) - L_n^{\text{PE}}(\hat{f}_\lambda) \leq \mathcal{O}(n^{-\frac{1}{2+\gamma}})$ with high probability.

5 COMPARISON WITH PREVIOUS WORKS

Previous LRE works based on Kernel Methods have focused on the offline setting, which assumes the availability of two data samples for training: $X = \{x_t \sim p\}_{t=1}^{n}$ and $X' = \{x_t \sim q\}_{t=1}^{n'}$. Within this context, the LRE problem translates to a convex optimization problem drawing upon the empirical approximation of a specific ϕ-divergence.

To scale to large n and n', most offline algorithms assume the likelihood-ratio can be approximated by a finite linear combination of the M elements of a fixed dictionary $D_M = \{K(x_m,\cdot)\}_{m=1}^{M}$, i.e. the approximation $\hat{f}$ should belong to $S = \text{span}(\{K(x_m,\cdot)\,|\,x_m \in D_M\})$ and hence can be written in the form:

$$\hat{f}(x) = \sum_{m=1}^{M} \hat{\theta}_m K(x_m, x) = K(D_M,\cdot)^\mathsf{T}\hat{\theta}. \tag{23}$$

Two notable methods for offline LRE are the *Relative Unconstrained Least-Squares Importance Fitting* (RULSIF) (Yamada et al., 2011) and the *Kullback-Leibler Importance Estimation Procedure* (KLIEP) (Sugiyama et al., 2007). To make the discussion about the properties of these methods more precise, let us introduce the following terms:

$$\begin{aligned} H &= \frac{1}{n}\sum_{x\in X} K(D_M,x)K(D_M,x)^\mathsf{T} \\ H' &= \frac{1}{n'}\sum_{x'\in X'} K(D_M,x')K(D_M,x')^\mathsf{T}, \\ h &= \sum_{x\in X}\frac{K(D_M,x)}{n}, \quad h' = \sum_{x'\in X'}\frac{K(D_M,x')}{n'}. \end{aligned} \tag{24}$$

RULSIF capitalizes over the χ^2-divergence to define a penalized empirical risk minimization problem to estimate r^α. This formulation along with Expr. 23 leads to an estimation that amounts to solving a linear system:

$$\begin{aligned} \hat{\theta}_{\text{RULSIF}} &= \underset{\theta\in\mathbb{R}^M}{\text{argmin}}\left(\frac{1-\alpha}{2}\right)\theta^\mathsf{T} H\theta + \left(\frac{\alpha}{2}\right)\theta^\mathsf{T} H'\theta \\ &\qquad - \theta^\mathsf{T} h' + \frac{\lambda}{2}\theta^\mathsf{T}\theta \\ &= \left(\frac{1-\alpha}{2}H + \frac{\alpha}{2}H' + \lambda I_M\right)^{-1} h', \end{aligned} \tag{25}$$

where $\lambda > 0$ is a fixed regularization constant.

On the other hand, KLIEP uses an empirical approximation of the KL-divergence to estimate the unregularized likelihood-ratio $r^{\alpha=0}(x) = \frac{q(x)}{p(x)}$. More precisely, the authors propose to solve the following constrained convex problem:

$$\begin{aligned} \hat{\theta}_{\text{KLIEP}} &= \underset{\theta\in\mathbb{R}^M}{\text{argmin}} - \sum_{x'\in X'} \log(K(D_M,x')^\mathsf{T}\theta) \\ &\text{such that } \theta^\mathsf{T} h = 1, \quad \theta_i \geq 0 \quad \forall i \in \{1,...,m\}. \end{aligned} \tag{26}$$

There are mainly two differences between RULSIF and KLIEP. First, KLIEP does not have a closed-form

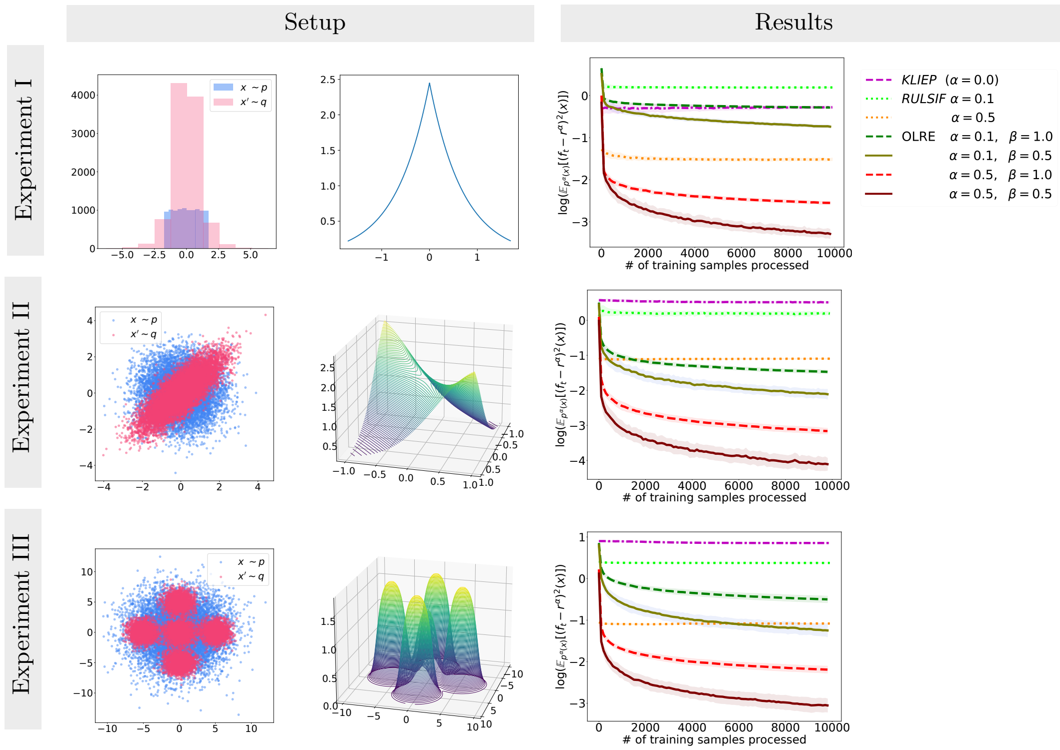

Figure 1: Each row presents results for one of the synthetic scenarios described in Sec. 6. The first column shows the generated samples from p and q, and the second column illustrates the real relative likelihood-ratio r^{α}. The third column compares the different algorithms in terms of the expected $\mathcal{L}^2_{p^{\alpha}}$-distance of the likelihood-ratio estimates f_t and the real r^{α}, as a function of the number of pairs of observations processed. The expected value $\mathbb{E}_{p^{\alpha}(y)}[(f_t - r^{\alpha})^2]$ is computed by averaging over 10000 independent pairs of observations that were not used during the training phase. The empirical convergence curve is the average of 100 experiment instances, and the error-bar indicates 1 standard deviation around the average performance. A safer comparison between two methods can be made when they use the same α-regularization, hence they optimize the same target likelihood-ratio functional.

solution, and the authors propose an iterative algorithm similar to the projected gradient descent. Second, the restrictions imposed in KLIEP aim at finding an estimator where the properties $r^{\alpha=0} > 0$ and $\mathbb{E}_{p(x)}[r^{\alpha=0}(x)] = 1$ are enforced.

RULSIF and KLIEP are not suitable for the online setting as observations are required to be known in advance. In order to compare their computational cost to that of OLRE, we assume we have waited for as long as we have accumulated T pairs of observations $\{(x_t, x'_t)\}_{t=1}^{T}$ and then we use all of them to estimate $\hat{\theta}_{\text{RULSIF}}$ and $\hat{\theta}_{\text{KLIEP}}$ (i.e. $M = 2T$). We compare the different methods in terms of the number of multiplications and additions (#MA) and kernel evaluations (#KE) they perform. In this setting, RULSIF has a complexity $\mathcal{O}(T^3)$ in #MA and $\mathcal{O}(T^2)$ in #KE. KLIEP needs $i_{\text{KLIEP}}(T)$ iterations before achieving convergence, the cost per iteration being $\mathcal{O}(T^2)$ in #MA and #KE. KLIEP's final computational cost is $\mathcal{O}(\#_{\text{KLIEP}}(T)T^2)$in #MA and $\mathcal{O}(T^2)$ in #KE. Notice that $i_{\text{KLIEP}}(T)$ is a function of T, since the error made by the optimization algorithm should not be bigger than the approximation error made by KLIEP that depends on the number of available observations. It has been shown that the error made by KLIEP decreases at rate $\mathcal{O}(T^{-\frac{1}{2}})$ (Sugiyama et al., 2008).

To reduce the dependency on T, it was suggested to prefix a dictionary with M elements selected uniformly at random from the observations. The computational gains of this approach are summarized in Tab. 1. Although this may seem an attractive approach, we will see in the experiments that it leads to a bias that does not disappear as the sample size increases. This result suggests that M should not be a fixed value, but rather a function that depends on the complexity of the LRE problem and the number of available observations. Nevertheless, this issue has not yet been addressed in the literature.

6 EXPERIMENTS

In this section, we carry out synthetic experiments to evaluate the performance of OLRE (Alg. 1), as well as its sensitivity to its hyperparameters. We compare OLRE variants against two existing offline approaches, more precisely RULSIF (Yamada et al., 2011, 2013)

Table 1: Compared LRE methods. For each method included in our experimental evaluation, we report its computational cost of approximating the likelihood-ratio when T pairs of incoming observations $\{(x_t, x'_t)\}_{t=1}^T$ are made available.

Method	Reference	Estimate	ϕ-divergence	Cost per iteration		Total cost	
				#MA	#KE	#MA	#KE
KLIEP	Sugiyama et al. (2007)	l.-r.	KL-divergence	$\mathcal{O}(T^2)$	$\mathcal{O}(T^2)$	$\mathcal{O}(i_{\text{KLIEP}}(T)T^2)$	$\mathcal{O}(T^2)$
				$\mathcal{O}(M^2)$	$\mathcal{O}(MT)$	$\mathcal{O}(i_{\text{KLIEP}}(T)MT)$	$\mathcal{O}(MT)$
RULSIF	Yamada et al. (2011)	relative l.-r.	χ^2-divergence	$\mathcal{O}(T^3)$	$\mathcal{O}(T^2)$	$\mathcal{O}(T^3)$	$\mathcal{O}(T^2)$
				$\mathcal{O}(M^3)$	$\mathcal{O}(MT)$	$\mathcal{O}(M^3)$	$\mathcal{O}(MT)$
OLRE	this work	relative l.-r.	χ^2-divergence	$\mathcal{O}(T)$	$\mathcal{O}(T)$	$\mathcal{O}(T^2)$	$\mathcal{O}(T^2)$

and KLIEP (Sugiyama et al., 2007, 2008). Since for large datasets the use of all the available observations becomes prohibitive for both methods, we follow the recommendation to select a random subset of basis functions for reducing their computational complexity; e.g. Sugiyama et al. (2007) take 100 basis functions associated to observations coming from q.

An important component of OLRE is the choice of the kernel function and its hyperparameters. We choose a Gaussian kernel, but other options are possible as mentioned in Sec. 4. To tune the kernel hyperparameters we perform cross-validation over the first $n = 100$ observations using RULSIF, which, as mentioned, has a closed-form and therefore allows for fast model selection. Following the results of Theorem 1, we let the learning rate and the penalization rate to depend on the smoothness of the parameters a, t_0, and β. We fix a at the lower-bound provided by Theorem 1, that is $a = 4$ and the lower-bound for t_0 is fixed as 100, which is equal to the number of observations used for identifying the hyperparameters at the beginning of the procedure. The user needs to provide only two parameters, α and β, which, according to Theorem 1, play an important role in OLRE's convergence. We report results with different values in order to show the sensibility of the method to parametrization.

We run experiments that approximate the likelihood-ratio between two pdfs p and q, using three setups:

- **Experiment I:** p is a uniform continuous distribution with zero mean and unit variance ($p = \mathcal{U}(-\sqrt{3}, \sqrt{3})$); q is a Laplace distribution with zero mean and unit variance.
- **Experiment II:** p is a bivariate Gaussian distribution with zero mean, and a covariance matrix equal to the identity matrix ($p = \mathcal{N}(\mathbf{0}_{2\times1}, I_{2\times2})$); q is a bivariate Gaussian distribution with zero mean and covariance matrix such that $\Sigma_{1,1} = \Sigma_{2,2} = 1$, $\Sigma_{1,2} = \frac{4}{5}$ ($q = \mathcal{N}(\mathbf{0}_{2\times1}, \Sigma)$).
- **Experiment III:** p is bivariate Gaussian distribution with mean vector μ and covariance matrix $\Sigma_1 = 10 \times I_{2\times2}$ ($p = \mathcal{N}(\mathbf{0}_{2\times1}, \Sigma_1)$), and q is a mixture of 5 bivariate Gaussian distributions with the same proportion and the same covariance matrix $\Sigma_2 = 5 \times I_{2\times2}$ and μ vectors: $\mu_1 = (0, 0), \mu_2 = (0, 5), \mu_3 = (0, -5), \mu_4 = (5, 0), \mu_5 = (-5, 0)$.

We compare the algorithms in approximating the relative likelihood-ratio r^α with respect to the norm $\|\cdot\|_{\mathcal{L}^2_{p^\alpha}}$, which is the real least-squared error $\mathbb{E}_{p^\alpha(y)}[(f_t - r^\alpha)^2]$, a quantity that we approximate by averaging over 10000 testing pairs of observations that were not used during training. For the offline setting, f_t stands for the estimated likelihood-ratio computed at each time from scratch by minimizing an empirical risk with respect to the first t pairs of training observations. For OLRE, f_t is the approximation to the solution of the functional minimization Problem 9 found via the functional stochastic gradient descent.

Fig. 1 reports our results that carry clear messages. The first thing to notice is that OLRE achieves substantially faster convergence rates when compared with available offline methods (when comparing for the same α value). Furthermore, we can see how KLIEP's and RULSIF's strategy of selecting a random dictionary introduces a bias to their performance over time. OLRE's behavior with respect to the hyperparameter α is well described by Theorem 1. Higher values of α lead to faster convergence. The value of β also impacts the performance of OLRE. Recall that a higher β value implies we assume r^α to be smoother with respect to the RKHS $\mathbb{H}$. Fig. 1 suggest also that higher values of β lead to a lower variance, but also a higher bias.

7 CONCLUSIONS

To the best of our knowledge, this is the first work to introduce and addresse the problem of online likelihood-ratio estimation (OLRE). We presented the homonymous non-parametric framework that processes a stream of pairs of observations coming from two pdfs. Our approach leads to an easy implementation that, contrary to the existing methods for the offline setting, does not require knowing the length of the stream in advance. Moreover, our theoretical results shed light on the limitations of previous convergence analyses and may motivate further work on studying the LRE problem with techniques used in functional optimization that can optimize directly the real risk.

ACKNOWLEDGMENTS

The authors acknowledge support from the Industrial Data Analytics and Machine Learning Chair hosted at ENS Paris-Saclay, Université Paris-Saclay.

References

Agrawal, R. and Horel, T. (2021). Optimal bounds between f-divergences and integral probability metrics. *Journal of Machine Learning Research*, 22(128):1–59.

Basseville, M. (2013). Divergence measures for statistical data processing — an annotated bibliography. *Signal Processing*, 93(4):621–633.

Bauschke, H. H. and Combettes, P. L. (2011). *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathematiques de la SMC. Springer, New York, NY.

Casella, G. and Berger, R. L. (2006). *Statistical Inference*. Thomson Press.

Csiszár, I. (1967). On topological properties of f-divergences. *Studia Scientiarum Mathematicarum Hungarica*, 2:329–339.

Dieuleveut, A. (2017). *Stochastic approximation in Hilbert spaces*. Theses, Université Paris sciences et lettres.

Dieuleveut, A. and Bach, F. (2016). Nonparametric stochastic approximation with large step-sizes. *The Annals of Statistics*, 44(4):1363 – 1399.

Fishman, G. S. (1996). *Monte Carlo*. Springer New York.

Kivinen, J., Smola, A., and Williamson, R. (2004). Online learning with kernels. *IEEE Transactions on Signal Processing*, 52(8):2165–2176.

Liese, F. and Vajda, I. (2006). On divergences and informations in statistics and information theory. *IEEE Trans. on Information Theory*, 52(10):4394–4412.

Liu, S., Yamada, M., Collier, N., and Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, 43:72–83.

Moustakides, G. V. and Basioti, K. (2019). Training neural networks for likelihood/density ratio estimation.

Neyman, J. and Pearson, E. S. (1933). IX. on the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694–706):289–337.

Nguyen, X., Wainwright, M. J., and Jordan, M. (2008). Estimating divergence functionals and the likelihood ratio by penalized convex risk minimization. In *Advances in Neural Information Processing Systems*.

Nguyen, X., Wainwright, M. J., and Jordan, M. I. (2010). Estimating divergence functionals and the likelihood ratio by convex risk minimization. *IEEE Transactions on Information Theory*, 56(11):5847–5861.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1–2):100–115.

Rubenstein, P., Bousquet, O., Djolonga, J., Riquelme, C., and Tolstikhin, I. O. (2019). Practical and consistent estimation of f-divergences. In *Advances in Neural Information Processing Systems*, volume 32.

Shiryaev, A. N. (1963). On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1):22–46.

Sugiyama, M., Nakajima, S., Kashima, H., Buenau, P., and Kawanabe, M. (2007). Direct importance estimation with model selection and its application to covariate shift adaptation. In *Advances in Neural Information Processing Systems*.

Sugiyama, M., Suzuki, T., and Kanamori, T. (2012). *Density Ratio Estimation in Machine Learning*. Cambridge University Press.

Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Bünau, P., and Kawanabe, M. (2008). Direct importance estimation for covariate shift adaptation. *Annals of the Institute of Statistical Mathematics*, 60:699–746.

Tarrès, P. and Yao, Y. (2014). Online learning as stochastic approximation of regularization paths: Optimality and almost-sure convergence. *IEEE Transactions on Information Theory*, 60(9):5716–5735.

Tartakovsky, A., Nikiforov, I., and Basseville, M. (2014). *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, CRC Press.

van de Geer, S. (2000). *Cambridge series in statistical and probabilistic mathematics: Empirical processes in M-estimation series number 6*. Cambridge University Press.

Xie, L., Zou, S., Xie, Y., and Veeravalli, V. V. (2021). Sequential (quickest) change detection: Classical results and new directions. *IEEE J. on Selected Areas in Information Theory*.

Yamada, M., Suzuki, T., Kanamori, T., Hachiya, H., and Sugiyama, M. (2011). Relative density-ratio estimation for robust distribution comparison. In *Advances in Neural Information Processing Systems*.

Yamada, M., Suzuki, T., Kanamori, T., Hachiya, H., and Sugiyama, M. (2013). Relative density-ratio estimation for robust distribution comparison. *Neural Computation*, 25(5):1324–1370.

Yao, Y. (2010). On complexity issues of online learning algorithms. *IEEE Transactions on Information Theory*, 56(12):6470–6481.

Zhang, Y. and Yang, Q. (2021). A survey on multi-task learning. *IEEE Trans. on Knowledge and Data Engineering*.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., and He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43–76.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. **Yes. (See Sec. 2 and Sec. 4)**
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. **Yes. (See Sec. 3)**
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. **Yes. (The code is provided as supplementary material.)**
2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. **Yes. (See Sec. 4)**
   (b) Complete proofs of all theoretical results. **Yes. The proofs appear in the supplementary materials.**
   (c) Clear explanations of any assumptions. **Yes. (See Sec. 4)**
3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results. **Yes. (The code is provided as supplemental material.)**
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). **Yes. (See Sec. 6)**
   (c) A clear definition of the specific measure or statistics and error bars. **Yes. (See Sec. 6)**
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). **Not Applicable, because the algorithm doesn't require a special computing infrastructure.**
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. **Yes. The models we used for comparisons are replicated by ourselves, cited, and available as supplementary material.**
   (b) The license information of the assets, if applicable. **Not Applicable.**
   (c) New assets either in the supplemental material or as a URL, if applicable. **Yes**
   (d) Information about consent from data providers/curators. **Not Applicable.**
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. **Not Applicable.**
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. **Not Applicable.**
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. **Not Applicable.**
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. **Not Applicable.**

A TECHNICAL RESULTS

This section contains the technical details of the results presented in Sec. 4. First, we introduce the necessary elements to define a linear operator equation in a Hilbert space, and then we present the regularized paths framework proposed in Tarrès and Yao (2014) for solving this kind of problems. After that, we detail how the likelihood-ratio estimation problem can be reformulated as a linear operator equation and its similarities with the regression problem in Hilbert spaces. Finally, we provide detailed proofs of Theorem 1 and 2.

A.1 Sequential stochastic approximations of regularization paths in Hilbert spaces

Tarrès and Yao (2014) consider the general case of minimizing a quadratic map defined over elements of a Hilbert space via stochastic approximation. Let us begin by denoting $SL(\mathbb{H})$ as the vector space of self-adjoint bounded linear operators on $\mathbb{H}$ endowed with the canonical norm:

$$\|A\| = \sup_{\|f\|_{\mathbb{H}} \leq 1} \|Af\|_{\mathbb{H}}, \qquad A \in SL(\mathbb{H}).$$

Notice that we have used the convention that Af denotes linear operator $A \in SL(\mathbb{H})$ applied to $f \in \mathbb{H}$. We will keep this notation for the rest of the section.

Let us denote by $\mathcal{X}$ and $\mathcal{Y}$ two topological spaces, and define $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ as their Cartesian product. We define a probability measure ρ on the Borel σ- algebra of $\mathcal{Z}$. Let $A : \mathcal{Z} \to SL(\mathbb{H})$ and $b : \mathcal{Z} \to \mathbb{H}$ be two random variables defined in terms of the space $\mathcal{Z}$ whose expected values are denoted by:

$$\mathbf{A} = \mathbb{E}_\rho[A], \qquad \mathbf{b} = \mathbb{E}_\rho[b].$$

The goal in Tarrès and Yao (2014) is to find $\mathbf{w} \in \mathbb{H}$ by solving the linear operator equation:

$$\mathbf{A}\mathbf{w} = \mathbf{b}.$$

where $\mathbf{A}$ and $\mathbf{b}$ are known, and $\mathbf{A}$ is a strictly positive operator with an unbounded inverse.

Alternatively, $\mathbf{w}$ can be defined as the solution to the quadratic optimization problem:

$$\operatorname*{argmin}_{f \in \mathbb{H}} Q(f) = \operatorname*{argmin}_{f \in \mathbb{H}} \frac{1}{2} \langle \mathbf{A}(f - \mathbf{w}), (f - \mathbf{w}) \rangle_{\mathbb{H}}. \tag{27}$$

The stochastic approximation approach proposed by Tarrès and Yao (2014) consists in defining a sequence of random variables $\{A_t\}_{t \in \mathbb{N}}$ and $\{b_t\}_{t \in \mathbb{N}}$ depending on the incoming observations $z_t = (x_t, y_t)$ such that the sequence $\{f_t\}_{t \in \mathbb{N}}$ generated by the iterative algorithm:

$$f_t(\cdot) = f_{t-1}(\cdot) - \eta_t \left(A_t(z_t) f_{t-1}(\cdot) - b_t(z_t)(\cdot) \right), \tag{28}$$

converges toward the solution of Problem 27.

In the same work, they study the required conditions to guarantee the convergence of Eq. 28 with respect to the norms $\|\cdot\|_{L^2_{p^\alpha}}$ and $\|\cdot\|_{\mathbb{H}}$. Among these conditions, the authors assume the random variables $\{A_t\}_{t \in \mathbb{N}}$ and $\{b_t\}_{t \in \mathbb{N}}$ are such that their expected values $\mathbf{A}_t = \mathbb{E}_\rho[A_t]$ and $\mathbf{b}_t = \mathbb{E}_\rho[b_t]$ satisfy $\mathbf{A}_t \to \mathbf{A}$ and $\mathbf{b}_t \to \mathbf{b}$, as $t \to \infty$, and each of the elements of the sequence $\{\mathbf{A}_t\}_{t \in \mathbb{N}}$ has a bounded inverse. Finally, the authors provide the required decreasing rate for the sequence of step-sizes $\{\eta_t\}_{t \in \mathbb{N}}$.

A.2 Application to the OLRE problem

The Online LRE described in Sec. 3 can be written in terms of the framework introduced in Tarrès and Yao (2014). In this context, $\mathcal{Z} = \mathcal{X} \times \mathcal{X} = \mathcal{X}^2$ and the associated probability measure is given by the joint pdf ρ with marginal pdfs p and q. The incoming data observations $z_t = (x_t, x'_t)$ are iid pairs such that $x_t \sim p$ and $x'_t \sim q$.

The random variables $A : \mathcal{Z} \to SL(\mathbb{H})$ and $b : \mathcal{Z} \to \mathbb{H}$ are defined based on the functional stochastic gradient:

$$A(x, x') = (1 - \alpha) \langle \cdot, K(x, \cdot) \rangle_{\mathbb{H}} K(x, \cdot) + \alpha \langle \cdot, K(x', \cdot) \rangle_{\mathbb{H}} K(x', \cdot) \qquad b(x, x') = K(x', \cdot). \tag{29}$$

Given the reproducing property of $\mathbb{H}$, we have that for $f \in \mathbb{H}$:

$$A(x, x')f = (1-\alpha)f(x)K(x,\cdot) + \alpha f(x')K(x',\cdot).$$

Under this configuration:

$$\begin{aligned}\mathbf{A} &= \mathbb{E}_{(p(x),q(x'))}[(1-\alpha)\langle \cdot, K(x,\cdot)\rangle_{\mathbb{H}} K(x,\cdot) + \alpha \langle \cdot, K(x',\cdot)\rangle_{\mathbb{H}} K(x',\cdot)] \\ &= \mathbb{E}_{p^\alpha(y)}[\langle \cdot, K(y,\cdot)\rangle_{\mathbb{H}} K(y,\cdot)] = \mathcal{L}_K,\end{aligned} \tag{30}$$

where the second equality is given by the linearity of the integral with respect to the mixture measure P^α and the definition of the covariance operator when restricted to elements of $\mathbb{H}$ (see Sec. 2.1).

$$\mathbf{b} = \mathbb{E}_{(p(x),q(x'))}[K(x',\cdot)] = \mathbb{E}_{p^\alpha(y)}[r^\alpha(y)K(y,\cdot)] = \mathcal{L}_K r^\alpha. \tag{31}$$

The second equality is given by the change of measure expression $\mathbb{E}_{q(x')}[g(x')] = \mathbb{E}_{p^\alpha(y)}[r^\alpha(y)g(y)]$, and the last one is due to the definition of the covariance operator and the hypothesis that $r^\alpha \in \mathbb{H}$ (see Eq. 19).

We can rewrite the LRE problem described in Eq. 9 as trying to minimize the quadratic function:

$$Q(f) = \langle \mathcal{L}_K(f - r^\alpha), f - r^\alpha\rangle_{\mathbb{H}} = \frac{1}{2}\mathbb{E}_{p^\alpha(y)}[(f - r^\alpha)^2(y)], \tag{32}$$

where the last equality is a consequence of property:

$$\langle f, \mathcal{L}_K(g)\rangle_{\mathbb{H}} = \mathbb{E}_{p^\alpha(y)}[f(y)g(y)] \quad \forall f, g \in \mathbb{H}. \tag{33}$$

The sequence of random variables $\{A_t\}_{t\in\mathbb{N}}$ and $\{b\}_{t\in\mathbb{N}}$ are given by the updates described in Alg. 1.

$$A_t = A((x_t, x'_t)) + \lambda_t I_{\mathbb{H}}; \quad b_t = K(x'_t, \cdot). \tag{34}$$

We can easily corroborate that $\mathbf{A}_t$ and $\mathbf{b}_t$ satisfy:

$$\begin{aligned}&\mathbf{A}_t = \mathcal{L}_K + \lambda_t I_{\mathbb{H}} \quad \text{and} \quad \mathbf{A}_t \to \mathbf{A} \quad \text{as} \quad \lambda_t \to 0; \\ &\mathbf{b}_t = \mathcal{L}_K r^\alpha.\end{aligned} \tag{35}$$

Moreover, by the properties of the covariance operator stated in Sec. 2.1, $\mathbf{A}_t$ has a bounded inverse.

After putting together these elements, we arrive to the stochastic approximation schema with the form:

$$\begin{aligned}f_t(\cdot) &= f_{t-1}(\cdot) - \eta_t\left[A_t f_{t-1}(\cdot) - b(x_t, x'_t)(\cdot)\right] \\ &= f_{t-1}(\cdot) - \eta_t\left[(1-\alpha)f_{t-1}(x_t)K(x_t,\cdot) + \alpha f_{t-1}(x'_t)K(x'_t,\cdot) + \lambda_t f_{t-1}(\cdot) - K(x'_t,\cdot)\right],\end{aligned} \tag{36}$$

which coincides with the functional stochastic gradient descent described in Eq. 16.

A term that will be important for studying the convergence of the online optimization schema is the solution to the regularized optimization problem:

$$f_{\lambda_t} = \operatorname*{argmin}_{f\in\mathbb{H}} \frac{1}{2}\langle \mathbf{A}_t(f - r^\alpha), f - r^\alpha\rangle_{\mathbb{H}} = \operatorname*{argmin}_{f\in\mathbb{H}} \frac{1}{2}\mathbb{E}_{p^\alpha(y)}[(f - r^\alpha)^2(y)] + \frac{\lambda_t}{2}\|f\|^2_{\mathbb{H}}. \tag{37}$$

In fact, f_{λ_t} can be written as:

$$f_{\lambda_t} = \mathbf{A}_t^{(-1)}\mathbf{b}_t = (\mathcal{L}_K + \lambda_t I_{\mathbb{H}})^{(-1)}\mathbf{b}_t. \tag{38}$$

A.3 Similarities between OLRE and Online Regression Problem

The framework described in Sec. A.1 was originally proposed to solve a regression problem in $\mathbb{H}$ as data observations arrive. In that context, $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ is the feature space and $\mathcal{Y} \subset \mathbb{R}$ represents noisy observations of the regression function to be approximated (f_ρ). ρ stands for the joint probability function of (x, y) whose marginal in the first entry is $\rho_{\mathcal{X}}$. The regression problem can be written as:

$$\min_{f\in\mathbb{H}} L^{\text{Reg}}(f) = \min_{f\in\mathbb{H}} \int_{\mathcal{X}\times\mathcal{Y}} (f(x) - y)^2 d\rho.$$

Given the previous problem, the random variables to be updated as (x_t, y_t) arrive take the form:

$$\begin{aligned} A^{\text{Reg}}(x,y) &= \langle \cdot, K(x,\cdot)\rangle_{\mathbb{H}} K(x,\cdot) & b^{\text{Reg}}(x,y) &= yK(x,\cdot) \\ \mathbf{A}^{\text{Reg}} &= \mathcal{L}_K^{\rho_x} & \mathbf{b}^{\text{Reg}} &= \mathcal{L}_K^{\rho_x} f_\rho \\ A_t^{\text{Reg}} &= A^{\text{Reg}}(x_t,y_t) + \lambda_t I_{\mathbb{H}} & b_t^{\text{Reg}} &= y_t K(x_t,\cdot) \\ \mathbf{A}_t^{\text{Reg}} &= \mathcal{L}_K^{\rho_x} + \lambda_t I_{\mathbb{H}} & \mathbf{b}_t^{\text{Reg}} &= \mathcal{L}_K^{\rho_x} f_\rho \end{aligned}$$

and $f_{\lambda_t}^{\text{Reg}} = (\mathcal{L}_K^{\rho_x} + \lambda_t I_{\mathbb{H}})^{(-1)} \mathbf{b}_t$.

As it can be seen, the main difference between Online Likelihood-Ratio Estimation and the Online Regression Problem is the definition of the random variables $\{A_t\}_{t\in\mathbb{N}}$ and $\{b_t\}_{t\in\mathbb{N}}$, while the expected values of these random variables as well as the regularized term f_{λ_t} take the same form. The covariance operator $\mathcal{L}_K^{\rho_x}$ translates to $\mathcal{L}_K$ defined in terms of the measure P^α and the regression function f_ρ to the relative likelihood-ratio r^α. These similarities facilitate the convergence analysis as we can reuse the results provided in Tarrès and Yao (2014) regarding the deterministic terms, and we only need to rework the terms involving the random variables $\{A_t\}_{t\in\mathbb{N}}$ and $\{b_t\}_{t\in\mathbb{N}}$.

A.4 Required elements for convergence analysis

The proof of Theorems 1 and 2 depends mainly on two iterative decompositions of the residuals between the solution to the approximation f_t and the solution to the regularization problem f_{λ_t}. A martingale decomposition will lead to convergence rates with respect to the norm $\|\cdot\|_{L^2_{p^\alpha}}$, while a reversed martingale decomposition will be useful when analyzing the convergence rates associated with the norm $\|\cdot\|_{\mathbb{H}}$.

For convenience, we denote by $\mathbb{E}[\cdot]$ the expected value with respect to the joint distribution $\mathbb{E}_{(p(x),q(x'))}[\cdot]$. We call Ξ_t the σ-algebra generated by the pairs of observations up to t, that is $\Xi_t = \sigma((x_1,x_1'),(x_2,x_2'),...,(x_t,x_t'))$. $\mathcal{B}_i$ denotes the sigma-algebra generated by the observation observed after i, $\mathcal{B}_i = \sigma((x_i,x_i'),(x_{i+1},x_{i+1}'),,...)$.

Martingale Decomposition. Let us denote by res_t the difference between the stochastic approximation f_t, obtained via function stochastic gradient descent, and f_{λ_t} the solution to the regularized Problem 37:

$$\begin{aligned} \text{res}_t &:= f_t - f_{\lambda_t} \\ &= f_{t-1} - \eta_t \left[A_t f_{t-1} - b_t\right] - f_{\lambda_t} \\ &= (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)(f_{t-1} - f_{\lambda_t}) + \eta_t \left[(\mathbf{A}_t - A_t) f_{t-1} + (b_t - \mathbf{A}_t f_{\lambda_t}))\right] \\ &= (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)(f_{t-1} - f_{\lambda_t}) + \eta_t \left[(\mathbf{A}_t - A_t) f_{t-1} + (b_t - \mathbf{b}_t)\right] \\ &= (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)(f_{t-1} - f_{\lambda_{t-1}}) - (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)(f_{\lambda_t} - f_{\lambda_{t-1}}) + \eta_t \left[(\mathbf{A}_t - A_t) f_{t-1} + (b_t - \mathbf{b}_t)\right] \\ &= (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)\,\text{res}_{t-1} - (I_{\mathbb{H}} - \eta_t \mathbf{A}_t)\Delta_t + \eta_t \epsilon_t, \end{aligned} \tag{39}$$

where we have used the iterative Alg. 28 and expression $\mathbf{b}_t = \mathbf{A}_t f_{\lambda_t}$ (Eq. 38). The term $\Delta_t := f_{\lambda_t} - f_{\lambda_{t-1}}$ denotes the difference between the solution of adjacent solutions to the regularized problem. The path $t \to f_{\lambda_t}$ is known as the *regularization path*. Finally, ϵ_t denotes the noise term:

$$\begin{aligned} \epsilon_t &:= (\mathbf{A}_t - A_t) f_{t-1} + (b_t - \mathbf{b}_t) \\ &= \left(\mathcal{L}_K - (1-\alpha)\langle \cdot, K(x_t,\cdot)\rangle_{\mathbb{H}} K(x_t,\cdot) + \alpha \langle \cdot, K(x_t',\cdot)\rangle_{\mathbb{H}} K(x_t',\cdot)\right) f_{t-1} + K(x_t',\cdot) - \mathcal{L}_K r^a \\ &= \mathbb{E}_{p^\alpha(y)}[f_{t-1}(y)K(y,\cdot)] - (1-\alpha) f_{t-1}(x_t) K(x_t,\cdot) - \alpha f_{t-1}(x_t') K(x_t',\cdot) \\ &+ K(x_t',\cdot) - \mathbb{E}_{p^\alpha(y)}[r^\alpha(y) K(y,\cdot)] \quad \text{(Eq. 19 and the first point of Expr. 2.)} \\ &= \mathbb{E}_{p^\alpha(y)}[f_{t-1}(y)K(y,\cdot)] - (1-\alpha) f_{t-1}(x_t) K(x_t,\cdot) - \alpha f_{t-1}(x_t') K(x_t',\cdot) + K(x_t',\cdot) - \mathbb{E}_{q(x')}[K(x',\cdot)]. \end{aligned} \tag{40}$$

If we iterate Expr. 39 up to $s \le t$:

$$res_t = \bar{\Pi}_{s+1}^t res_s - \sum_{j=s+1}^{t} \bar{\Pi}_j^t \Delta_j + \sum_{j=s+1}^{t} \eta_j \bar{\Pi}_{j+1}^t \epsilon_j, \tag{41}$$

$$\bar{\Pi}_j^t = \begin{cases} \prod_{i=j}^t (I_{\mathbb{H}} - \eta_i \mathbf{A}_i) & \text{, if } j \le t; \\ I_{\mathbb{H}}, & \text{otherwise.} \end{cases} \tag{42}$$

From Eq. 40 and the independence of incoming observations, it is easy to verify that the process $\{\eta_j \bar{\Pi}^t_{j+1}\epsilon_j\}_{j\in\mathbb{N}}$ defines a martingale difference with respect to the filtration $\{\Xi_t\}_{t\in\mathbb{N}}$. The decomposition of Eq. 41 was first proposed in Yao (2010). The proof of Theorem 2 consists in finding an upperbound for the norm of each of the three terms in Eq. 41 and the residual difference $f_{\lambda_t} - r^\alpha$.

Reversed Martingale Decomposition.

Let us define the random operator in terms of the sample $((x_1, x'_1), (x_2, x'_2), ..., (x_n, x'_n))$ indexed by $j, t \in \mathbb{N}$:

$$\Pi^t_j(\{(x_i, x'_i)\}_{i\in\mathbb{N}}) = \begin{cases} \prod_{i=j}^t (I_{\mathbb{H}} - \eta_i A_i(x_i, x'_i)), & \text{if } j \le t; \\ I_{\mathbb{H}}, & \text{otherwise.} \end{cases} \tag{43}$$

We recover an alternative decomposition for the residual res_t:

$$\begin{aligned} res_t &= f_t - f_{\lambda_t} \\ &= f_{t-1} - f_{\lambda_t} - \eta_t(A_t f_{t-1} - b_t) \\ &= (I_{\mathbb{H}} - \eta_t A_t)(f_{t-1} - f_{\lambda_{t-1}}) - (I_{\mathbb{H}} - \eta_t A_t)(f_{\lambda_t} - f_{\lambda_{t-1}}) - \eta_t(A_t f_{\lambda_t} - b_t) \\ &= (I_{\mathbb{H}} - \eta_t A_t) res_{t-1} - (I_{\mathbb{H}} - \eta_t A_t)\Delta_t - \eta_t(A_t f_{\lambda_t} - b_t). \end{aligned}$$

By iterating the last expression for $s \le t$, we recover the following equality:

$$res_t = \Pi^t_{s+1} res_s - \sum_{j=s+1}^t \Pi^t_j \Delta_j - \sum_{j=s+1}^t \eta_j \Pi^t_{j+1}(A_j f_{\lambda_j} - b_j). \tag{44}$$

This decomposition was first introduced in Tarrès and Yao (2014).

Let us show that $\{\Pi^t_{j+1}(A_j f_{\lambda_j} - b_j)\}_{j\in\mathbb{N}}$ is a reversed martingale difference with respect to $\{\mathcal{B}_j\}_{\{j\in\mathbb{N}\}}$.

Definition 1. *Let $\{\mathcal{B}_i\}_{\{i\in\mathbb{N}\}}$ be a decreasing sequence of sub-σ-fields of $\mathcal{A}$ in the probability space $(\mathcal{Z}, \mathcal{A}, \rho)$. A sequence $\{\zeta_i\}_{i\in\mathbb{N}}$ integrable real random variables is called a reversed martingale difference if:*

1. *The real random variable ζ_i is $\mathcal{B}_i$-measurable for all $i \in \mathbb{N}$,*
2. *$\mathbb{E}[\zeta_i \,|\, \mathcal{B}_{i+1}] = 0$ for all $i \in \mathbb{N}$*

The term $\eta_j \Pi^t_{j+1}(A_j f_{\lambda_j} - b_j)$ defines a reversed martingale with respect to the sequence $\mathcal{B}_j = \sigma((x_j, x'_j), ..., (x_t, x'_t), ...)$. From its definition $\eta_j \Pi^t_{j+1}(A_j f_{\lambda_j} - b_j)$ is $\mathcal{B}_j$ measurable, moreover given the independence of the observations we have:

$$\begin{aligned} \mathbb{E}\left[\eta_j \Pi^t_{j+1}(A_j f_{\lambda_j} - b_j) \,|\, \mathcal{B}_{j+1}\right] &= \eta_j \Pi^t_{j+1} \mathbb{E}\left[A_j f_{\lambda_j} - b_j \,|\, \mathcal{B}_{j+1}\right] \\ &= \eta_j \Pi^t_{j+1}\left(\mathbf{A}_j f_{\lambda_j} - \mathbf{b}_j\right) \quad (\text{By the independence hypothesis}) \\ &= 0 \quad (\text{Eq. } 38). \end{aligned}$$

A.5 Convergence in $L^2_{p^\alpha}$

The proof of Theorem 1 mimics the proof of Theorem C in Tarrès and Yao (2014). This theorem is stated in Online Linear Regression and built upon decomposition of Eq. 41. As explained in Sec. A.3, the regression problem is similar to OLRE with the main difference being the operators A_t and b_t. This difference requires us to rework the bounds depending on the random processes $\{A_t\}_{t\in\mathbb{N}}$ and $\{b_t\}_{t\in\mathbb{N}}$.

Let us start by analyzing the $L^2_{p^\alpha}-$norm of the residuals:

$$\begin{aligned} \|f_t - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} &\le \|f_{\lambda_t} - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} + \|f_t - f_{\lambda_t}\|_{\mathcal{L}^2_{p^\alpha}} \\ &\le \|f_{\lambda_t} - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} + \left\|\bar{\Pi}^t_1 res_0\right\|_{\mathcal{L}^2_{p^\alpha}} + \left\|\sum_{j=1}^t \bar{\Pi}^t_j \Delta_j\right\|_{\mathcal{L}^2_{p^\alpha}} + \left\|\sum_{j=1}^t \eta_j \bar{\Pi}^t_{j+1}\epsilon_j\right\|_{\mathcal{L}^2_{p^\alpha}} \\ &= \mathcal{E}_{\text{init}}(t) + \mathcal{E}_{\text{drift}}(t) + \mathcal{E}_{\text{approx}} + \mathcal{E}_{\text{sample}}(t), \end{aligned} \tag{45}$$

where the second line comes from the martingale decomposition of Eq. 41 applied to s=0. Each of the error terms in Eq. 45 is defined as:

$$\begin{aligned} &\mathcal{E}_{\text{init}}(t) := \left\|\bar{\Pi}_1^t res_0\right\|_{\mathcal{L}^2_{p^\alpha}} && \mathcal{E}_{\text{approx}}(t) := \|f_{\lambda_t} - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}}, \\ &\mathcal{E}_{\text{drift}}(t) := \left\|\sum_{j=1}^t \bar{\Pi}_j^t \Delta_j\right\|_{\mathcal{L}^2_{p^\alpha}} && \mathcal{E}_{\text{sample}}(t) := \left\|\sum_{j=1}^t \eta_j \bar{\Pi}_{j+1}^t \epsilon_j\right\|_{\mathcal{L}^2_{p^\alpha}} \end{aligned} \tag{46}$$

The first three terms have the same behavior in the OLRE and Regression Problem, as they depend solely on equivalent deterministic terms, meaning we can reuse the upper bounds from Tarrès and Yao (2014). For completeness of our presentation, we restate these results. The last term differs and an upperbound is derived in Theorem 6.

For the following statements $\bar{t} = t + t_0$, and $t_0 > 0$ will be a given integer, a, b are two positive constants, β is the parameter related to the smoothness of a function in $\mathbb{H}$ as it was explained in Sec. 2.1 and α is the regularized parameter of the relative likelihood-ratio function r^α.

Theorem 3. ***(Theorem VI.1 in Tarrès and Yao (2014))*** *Let $t_0^\theta \geq a(C^2 + b)$. Then for all $t \in \mathbb{N}$.*

$$\mathcal{E}_{\text{init}}(t) \leq \frac{1}{\alpha}\left(\frac{t_0+1}{\bar{t}}\right)^{ab} \leq B_1 \bar{t}^{-ab}, \tag{47}$$

where $B_1 = \frac{(t_0+1)^{ab}}{\alpha}$.

Theorem 4. ***(Theorem VI.2 in Tarrès and Yao (2014))*** *For $\beta \in (0, 1]$ and $\mathcal{L}_K^{(-\beta)} r^\alpha \in \mathcal{L}^2_{p^\alpha}$,*

$$\mathcal{E}_{\text{approx}}(t) \leq \frac{b^\beta \bar{t}^{(-\beta(1-\theta))} \left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}}{\beta} \leq B_2 b^\beta \bar{t}^{(-\beta(1-\theta))}, \tag{48}$$

where $B_2 = \frac{\left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}}{\beta}$.

Theorem 5. ***(Theorem VI.3 in Tarrès and Yao (2014))*** *Assume $t_0^\theta = [a(C^2 + b) \vee 1]$. Then, if $\beta \in (0, 1]$ and $\mathcal{L}_K^{(-\beta)} r^\alpha \in \mathcal{L}^2_{p^\alpha}$:*

$$\mathcal{E}_{\text{drift}}(t) = \begin{cases} B_3 b^\beta \bar{t}^{-\beta(1-\theta)} & \text{if } ab > \beta(1-\theta); \\ B_3 b^\beta \bar{t}^{-ab} & \text{if } ab < \beta(1-\theta), \end{cases} \tag{49}$$

where $B_3 = \frac{4(1-\theta)}{|ab - \beta(1-\theta)|} \left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}$.

Theorem 6. *Assume that $\mathcal{L}_K^{(-\beta)}(r^\alpha) \in \mathcal{L}^2_{p^\alpha}$ for some $\beta \in [\frac{1}{2}, 1]$, $\theta \in [\frac{1}{2}, \frac{2}{3}]$, $ab = 1$, $a \geq 4$ and $t_0^\theta \geq 2 + 4C^2 a$. Then, for all $t \in \mathbb{N}$, with probability at least $1 - \delta$:*

$$\mathcal{E}_{\text{sample}}(t) \leq \frac{\sqrt{a} B_4}{\bar{t}^{\frac{\theta}{2}}} \log\left(\frac{2}{\delta}\right) + \left[B_5 a^{\frac{5}{2}} + B_6 a^{\frac{7}{2}} \sqrt{\log \bar{t}}\right] \frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}, \tag{50}$$

where:

$$B_4 = \frac{16C}{\alpha} \quad B_5 = \frac{32C^3}{\alpha} \quad B_6 = \frac{8C^3(10C+3)}{\alpha}.$$

The proof of the last theorem is given in Sec. A.7.

Proof of Theorem 1

Proof. By putting together the conditions stated in the statement of Theorem 1 and fixing $\theta = \frac{2\beta}{2\beta+1}$, $a \geq 4$,

$b \leq \frac{1}{4}$ such that $ab = 1$ and $t_0^\theta \geq 4aC^2 + 2$ we can verify that the requirements of Theorems 3-6 are satisfied:

$$\begin{aligned}\|f_t - r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} &\leq \mathcal{E}_{\text{init}}(t) + \mathcal{E}_{\text{approx}}(t) + \mathcal{E}_{\text{drift}}(t) + \mathcal{E}_{\text{sample}}(t) \\ &\leq \frac{B_1}{\bar{t}} + \left((B_2 + B_3)a^{-\beta} + \sqrt{a}B_4 \log\left(\frac{2}{\delta}\right)\right)\left(\frac{1}{\bar{t}}\right)^{\frac{\beta}{2\beta+1}} + \left(B_5 a^{\frac{5}{2}} + B_6 a^{\frac{7}{2}}\sqrt{\log(\bar{t})}\right)\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{4\beta-1}{4\beta+2}}} \\ &= \frac{C_1}{\bar{t}} + \left(C_2 a^{-r} + C_3\sqrt{a}\log\left(\frac{2}{\delta}\right)\right)\left(\frac{1}{\bar{t}}\right)^{\frac{\beta}{2\beta+1}} + \left(C_4 a^{\frac{5}{2}} + C_5 a^{\frac{7}{2}}\sqrt{\log(\bar{t})}\right)\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{4\beta-1}{4\beta+2}}}\end{aligned} \tag{51}$$

where $C_1 = \frac{2t_0}{\alpha} \geq \frac{t_0+1}{\alpha}$, $C_2 = B_2 + B_3 = \frac{5\beta+1}{\beta(1+\beta)}\left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}$, $C_3 = B_4$, $C_4 = B_5$ and $C_5 = B_6$. □

A.6 Convergence in $\mathbb{H}$

The study of the norm in $\mathbb{H}$ has a lot of similarities with the analysis of $L^2_{p^\alpha}$, starting with the decomposition of the norm into four terms that will be upper-bounded independently:

$$\begin{aligned}\|f_t - r^\alpha\|_{\mathbb{H}} &\leq \|f_t - f_{\lambda_t}\|_{\mathbb{H}} + \|f_{\lambda_t} - r^\alpha\|_{\mathbb{H}} \\ &\leq \|f_{\lambda_t} - r^\alpha\|_{\mathbb{H}} + \left\|\Pi_1^t res_0\right\|_{\mathbb{H}} + \left\|\sum_{j=1}^t \Pi_j^t \Delta_j\right\|_{\mathbb{H}} + \left\|\sum_{j=1}^t \eta_j \Pi_{j+1}^t (A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}} \quad \text{(Eq. 44)} \\ &= \mathcal{E}'_{\text{init}}(t) + \mathcal{E}'_{\text{drift}}(t) + \mathcal{E}'_{\text{approx}}(t) + \mathcal{E}'_{\text{sample}}(t),\end{aligned} \tag{52}$$

where,

$$\begin{aligned}\mathcal{E}'_{\text{init}}(t) &:= \left\|\Pi_j^t res_0\right\|_{\mathbb{H}} & \mathcal{E}'_{\text{approx}}(t) &:= \|f_{\lambda_t} - r^\alpha\|_{\mathbb{H}}, \\ \mathcal{E}'_{\text{drift}}(t) &:= \left\|\sum_{j=1}^t \Pi_j^t \Delta_j\right\|_{\mathbb{H}} & \mathcal{E}'_{\text{sample}}(t) &:= \left\|\sum_{j=1}^t \eta_j \Pi_{j+1}^t (A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}}\end{aligned} \tag{53}$$

As in the previous case, we will start by restating the required elements from Tarrès and Yao (2014) to upper-bound the deterministic terms of Eq. 52.

Theorem 7. ***(Theorem V.1 in Tarrès and Yao (2014))*** *Let $t_0^\theta \geq a(C^2 + b)$. Then, for all $t \in \mathbb{N}$,*

$$\mathcal{E}'_{init}(t) \leq B'_1 \bar{t}^{-ab}, \tag{54}$$

where $B'_1 = (t_0 + 1)^{ab}\|f_{\lambda_0}\|_{\mathbb{H}}$.

Theorem 8. ***(Theorem V.2 in Tarrès and Yao (2014))*** *For $\beta \in (\frac{1}{2}, \frac{3}{2}]$ and $\mathcal{L}_K^{(-\beta)} r^\alpha \in \mathcal{L}^2_{p^\alpha}$*

$$\mathcal{E}'_{\text{approx}}(t) \leq B'_2 b^{\beta-\frac{1}{2}} \bar{t}^{-(\beta-\frac{1}{2})(1-\theta)}, \tag{55}$$

where $B'_2 = (\beta - \frac{1}{2})^{-1}\left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}$.

Theorem 9. ***(Theorem V.3 in Tarrès and Yao (2014))*** *Let $t_0^\theta \geq \max(a(C^2 + b), 1)$. Then, for $\beta \in (\frac{1}{2}, \frac{3}{2}]$ and $\mathcal{L}_K^{(-\beta)} r^\alpha \in \mathcal{L}^2_{p^\alpha}$,*

$$\mathcal{E}'_{\text{drift}}(t) = \begin{cases} B'_3 b^{\beta-\frac{1}{2}} \bar{t}^{-(\beta-\frac{1}{2})(1-\theta)} & \text{if } ab > (\beta - \frac{1}{2})(1-\theta); \\ B'_3 b^{\beta-\frac{1}{2}} \bar{t}^{-ab} & \text{if } ab < (\beta - \frac{1}{2})(1-\theta), \end{cases} \tag{56}$$

where $B'_3 = \frac{4(1-\theta)}{|ab-(\beta-\frac{1}{2})(1-\theta)|}\left\|\mathcal{L}_K^{(-\beta)} r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}$.

Theorem 10. *Assume that $t_0^\theta \geq \min\{a(C^2 + b), b, 1\}$, $t_0^{1-\theta} \geq b$ and $ab \neq \theta - \frac{1}{2}$ or $ab \neq \frac{3\theta-1}{2}$. Then, with probability at least $1 - \delta$ ($\delta \in (0, 1)$),*

$$\mathcal{E}'_{\text{sample}}(t) \leq ab^{-\frac{1}{2}} B'_4 \bar{t}^{-\left(ab \wedge \frac{3\theta-1}{2}\right)} + B'_5 a \bar{t}^{-\left(ab \wedge (\theta - \frac{1}{2})\right)}, \tag{57}$$

where

$$B'_4 = \frac{2e}{3}\left(\frac{C+1}{\alpha} + C\right)\log\left(\frac{2}{\delta}\right) \quad B'_5 = 2\sqrt{\frac{1}{|ab - \theta + \frac{1}{2}|}} eC \log\left(\frac{2}{\delta}\right). \tag{58}$$

The proof of Theorem 10 is provided in Appendix A.7 and it capitalizes over the properties of the operator A_t and Lemma 9.

Proof of Theorem 2.

Proof. Let us fix $\theta = \frac{2\beta}{2\beta+1}$, $a \geq 1$,$b \leq 1$ such that $ab = 1$ and we assume $t_0^\theta \geq aC^2 + 1$. These constants imply that $\theta < 1$ and $t_0 \geq 1$ which means $t_0^{1-\theta} \geq b$, $ab = 1 \neq \theta - \frac{1}{2}$, $ab \neq \frac{(3\theta-1)}{2}$ and $(\beta - \frac{1}{2})(1-\theta) = \frac{2\beta-1}{4\beta+2} \leq 1$. Then, the hypothesis of Theorems 7-10 are satisfied and we can conclude:

$$\begin{aligned}\|f_t - r^\alpha\|_{\mathbb{H}} &\leq B_1'\bar{t}^{-ab} + B_2' b^{\beta-\frac{1}{2}}\bar{t}^{-(\beta-\frac{1}{2})(1-\theta)} + B_3' b^{\beta-\frac{1}{2}}\bar{t}^{-(\beta-\frac{1}{2})(1-\theta)} + ab^{-\frac{1}{2}}B_4'\bar{t}^{-\left(ab\wedge\frac{3\theta-1}{2}\right)} + B_5' a\bar{t}^{-\left(ab\wedge(\theta-\frac{1}{2})\right)} \\ &\leq B_1'\bar{t}^{-ab} + B_2' a^{\frac{1}{2}-\beta}\bar{t}^{-\frac{2\beta-1}{4\beta+2}} + B_3' a^{\frac{1}{2}-\beta}\bar{t}^{-\frac{2\beta-1}{4\beta+2}} + aa^{\frac{1}{2}}B_4'\bar{t}^{-\theta}\bar{t}^{-\frac{2\beta-1}{4\beta+2}} + B_5' a\bar{t}^{-\frac{2\beta-1}{4\beta+2}} \\ &= C_1'\bar{t}^{-ab} + \left[C_2' a^{\frac{1}{2}-\beta} + C_3' a \log\left(\frac{2}{\delta}\right)\right]\bar{t}^{-\frac{2\beta-1}{4\beta+2}},\end{aligned} \tag{59}$$

where:

$$\begin{aligned}B_1' &= (t_0+1)\|f_{\lambda_0}\|_{\mathbb{H}} \leq \frac{t_0+1}{\alpha\sqrt{\lambda_0}} \qquad (\text{Eq. } 121) \\ &\leq \frac{2t_0(t_0)^{\frac{\theta-1}{2}}}{\alpha b^{\frac{1}{2}}} = \frac{2t_0^{\frac{4\beta+1}{4\beta+2}}}{\alpha b^{\frac{1}{2}}} := C_1' \\ C_2' &:= B_2' + B_3' = \left((\beta-\frac{1}{2})^{-1} + \frac{4(1-\theta)}{\left|ab-(\beta-\frac{1}{2})(1-\theta)\right|}\right)\left\|\mathcal{L}_K^{(-\beta)}r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}} \\ &= \left(\frac{2}{2\beta-1} + \frac{8}{2\beta+3}\right)\left\|\mathcal{L}_K^{(-\beta)}r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}} = \frac{20\beta-2}{(2\beta-1)(2\beta+3)}\left\|\mathcal{L}_K^{(-\beta)}r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}}\end{aligned}$$

$$\begin{aligned}\sqrt{a}B_4'\bar{t}_0^{-\theta} + B_5' &\leq \frac{2e}{3}\left(\frac{1}{\alpha} + \frac{1}{\alpha C} + 1\right)\log\left(\frac{2}{\delta}\right) + 2\sqrt{\frac{1}{\left|ab-\theta+\frac{1}{2}\right|}}eC\log\left(\frac{2}{\delta}\right) \qquad \left(\text{The fact } t_0^\theta \geq aC^2+1 \text{ implies } \sqrt{a}t_0^{-\frac{\theta}{2}} \leq \frac{1}{C}\right) \\ &\leq e\log\left(\frac{2}{\delta}\right)\left(3C + \frac{2}{3}\left(\frac{C+1}{C\alpha}+1\right)\right) \leq 2\log\left(\frac{2}{\delta}\right)\left(\frac{9C^2\alpha + C(2+\alpha)+2}{3C\alpha}\right) \\ &\leq 2\log\left(\frac{2}{\delta}\right)\frac{(3C+\sqrt{2})^2}{3C\alpha} \leq 6\log\left(\frac{2}{\delta}\right)\frac{(C+1)^2}{C\alpha} \leq C_3'\log\left(\frac{2}{\delta}\right).\end{aligned}$$

□

A.7 Upperbounds on the noise terms

The goal of this section is to detail the upper-bounds on the noise terms presented in Theorems 6 and 10.

For the remainder of the discussion we will fix the step-size and the regularization constant sequence as:

$$\eta_t = \frac{a}{(t+t_0)^\theta} = \frac{a}{\bar{t}^\theta} \qquad \lambda_t = \frac{b}{(t+t_0)^{1-\theta}} = \frac{b}{\bar{t}^{1-\theta}}, \qquad \text{for some} \qquad \theta \in [0,1], t_0 > 0. \tag{60}$$

Let us begin with the definition of the following stochastic processes:

$$\begin{aligned} L_t &= \langle\cdot, K(x_t,\cdot)\rangle_{\mathbb{H}}K(x_t,\cdot), & \mathbf{L} &= \mathbb{E}\left[L_t \,|\, \Xi_{t-1}\right] \\ R_t &= \langle\cdot, K(x_t',\cdot)\rangle_{\mathbb{H}}K(x_t',\cdot), & \mathbf{R} &= \mathbb{E}\left[R_t \,|\, \Xi_{t-1}\right] \\ LR_t &= (1-\alpha)\langle\cdot, K(x_t,\cdot)\rangle_{\mathbb{H}}K(x_t,\cdot) + \alpha\langle\cdot, K(x_t',\cdot)\rangle_{\mathbb{H}}K(x_t',\cdot) & \mathbf{LR} &= \mathbb{E}\left[(1-\alpha)L_t + \alpha R_t \,|\, \Xi_{t-1}\right].\end{aligned} \tag{61}$$

Notice that the operator $\mathbf{LR}$ coincides with the covariance operator for $f \in \Xi_{t-1}$ (Expr. 19):

$$\mathbf{LR}(f) = \mathbb{E}\left[(1-\alpha)f(x_t)K(x_t,\cdot) + \alpha f(x_t')K(x_t',\cdot) \,|\, \Xi_{t-1}\right] = \mathbb{E}_{p^\alpha(y)}[f(y)K(y,\cdot)] = \mathcal{L}_K(f) \tag{62}$$

Upper-bound for $\mathcal{E}_{sample}(t) = \left\|\sum_{j=s+1}^{t}\eta_j\bar{\Pi}_{j+1}^t\epsilon_j\right\|_{\mathcal{L}^2_{p^\alpha}}$.

We can rewrite the residual between the approximation at time f_t and the relative likelihood-ratio r^α in terms of the stochastic processes defined in Eq. 61:

$$\begin{aligned} f_t - r^\alpha &= \left[I_{\mathbb{H}} - \eta_t\left((1-\alpha)L_t + \alpha R_t + \lambda_t I_{\mathbb{H}}\right)\right](f_{t-1}) + \eta_t K(x'_t,\cdot) - r^\alpha \\ &= \left[I_{\mathbb{H}} - \eta_t\left((1-\alpha)L_t + \alpha R_t + \lambda_t I_{\mathbb{H}}\right)\right](f_{t-1} - r^\alpha) + \eta_t\left(K(x'_t,\cdot) - \left[(1-\alpha)L_t + \alpha R_t\right] r^\alpha\right) - \eta_t\lambda_t r^\alpha, \end{aligned} \tag{63}$$

Let us define the sequences $\{g_t\}_{t\in\mathbb{N}}$, $\{h_t\}_{t\in\mathbb{N}}$:

$$g_0 = -r^\alpha \qquad h_0 = 0,$$

and

$$\begin{aligned} g_t &= \left[I_{\mathbb{H}} - \eta_t\left(\mathbf{LR} + \lambda_t I_{\mathbb{H}}\right)\right] g_{t-1} - \eta_t\lambda_t r^\alpha \\ h_t &= \left[I_{\mathbb{H}} - \eta_t\left(LR_t + \lambda_t I_{\mathbb{H}}\right)\right] h_{t-1} + \eta_t\left(K(x'_t,\cdot) - LR_t r^\alpha\right) + \eta_t\left[\mathbf{LR} - LR_t\right] g_{t-1}. \end{aligned} \tag{64}$$

By induction over Expr. 64, we can verify:

$$f_t - r^\alpha = g_t + h_t. \tag{65}$$

Notice that $\{g_t\}_{t\in\mathbb{N}}$ is a deterministic sequence, while $\{h_t\}_{t\in\mathbb{N}}$ is a random one.

We can use the aforementioned variables to upperbound the Hilbert norm of the noise term as follows:

$$\begin{aligned} \mathbb{E}\left[\|\epsilon_t\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] &= \mathbb{E}\left[\|K(x'_t,\cdot) - (1-\alpha)L_t f_{t-1} - \alpha R_t f_{t-1} - (\mathcal{L}_K r^\alpha - \mathcal{L}_K f_{t-1})\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] && (\text{Eq. } 40) \\ &\leq \mathbb{E}\left[\|K(x'_t,\cdot) - (1-\alpha)L_t f_{t-1} - \alpha R_t f_{t-1}\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] && (\text{After developing the norm and using Eq. } 31 \text{ and Eq. } 62) \\ &= \mathbb{E}\left[\|K(x'_t,\cdot) - (1-\alpha)L_t\left(r^\alpha + g_{t-1} + h_{t-1}\right) - \alpha R_t\left(r^\alpha + g_{t-1} + h_{t-1}\right)\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] && (\text{Eq. } 65) \\ &\leq 3\Bigg[\mathbb{E}\left[\|K(x'_t,\cdot) - \left[(1-\alpha)L_t + \alpha R_t\right] r^\alpha\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] + \mathbb{E}\left[\|\left[(1-\alpha)L_t + \alpha R_t\right] g_{t-1}\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right] \\ &\quad + \mathbb{E}\left[\|\left[(1-\alpha)L_t + \alpha R_t\right] h_{t-1}\|^2_{\mathbb{H}} \mid \Xi_{t-1}\right]\Bigg] && (\text{Inequality } 2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|^2_{\mathbb{H}} + \|b\|^2_{\mathbb{H}}). \end{aligned} \tag{66}$$

For the rest of this section, we will focus on providing an upperbound for each of the terms in Eq. 66.

Let us start by analyzing the deterministic sequence in $\{g_t\}_{t\in\mathbb{N}}$. We rewrite the following inequality shown in Tarrès and Yao (2014) (Lemma B.3):

Lemma 2. *Assume $t_0^\theta \geq a(C^2 + b)$. Then, for all $t \in \mathbb{N}$,*

1. $\|g_t\|_{\mathcal{L}^2_{p^\alpha}} \leq \frac{1}{\alpha}$
2. $\|g_t + r^\alpha\|_{\mathbb{H}} \leq \frac{3}{\alpha\sqrt{\lambda_t}}$.

As a consequence of this result we can easily verify the following inequality:

Lemma 3. *g_{t-1} satisfies the following inequalities:*

$$\|\mathbf{LR} g_{t-1}\|^2_{\mathbb{H}} \leq \frac{C^2}{\alpha^2}. \tag{67}$$

Proof. As g_{t-1} is $\Xi_{t-1}-$ measurable we have that $\mathbf{LR}(g_{t-1}) = \mathcal{L}_K g_{t-1}$ (see 62),then:

$$\begin{aligned} \|\mathbf{LR}(g_{t-1})\|^2_{\mathbb{H}} = \langle \mathcal{L}_K(g_{t-1}), \mathcal{L}_K(g_{t-1})\rangle_{\mathbb{H}} &= \mathbb{E}_{p^\alpha(y)}[\mathcal{L}_K(g_{t-1})(y) g_{t-1}(y)] && (\text{Eq. } 33) \\ &= \mathbb{E}_{p^\alpha(y)}[\int K(x,y) g_{t-1}(x) g_{t-1}(y) dP^\alpha(x)] && (\text{Eq. } 19) \\ &\leq C^2\left(\mathbb{E}_{p^\alpha(y)}[g_{t-1}(y)]\right)^2 && (\text{Assumption } 2) \\ &\leq C^2 \mathbb{E}_{p^\alpha(y)}[g^2_{t-1}(y)] && (\text{Jensen's inequality}) \\ &\leq \frac{C^2}{\alpha^2} && (\text{Lemma } 2). \end{aligned}$$

□

Now, let us continue with the random sequence $\{h_t\}_{t\in\mathbb{N}}$. We will start by defining the following operators for $t \in \mathbb{N}$ and $M_t \in \mathbb{R}_+ \cup \{+\infty\}$, which will allow us to upper-bound the norm of h_t with respect to a random variable with a bounded variance:

$$
\begin{aligned}
\overline{L}_t &= \mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\}L_t & \overline{\mathbf{L}} &= \mathbb{E}[\overline{L}_t \,|\, \Xi_{t-1}] \\
\underline{L}_t &= \mathbb{1}\{|h_{t-1}(x_t)| \geq M_t\}L_t & \underline{\mathbf{L}} &= \mathbb{E}[\underline{L}_t \,|\, \Xi_{t-1}] \\
\overline{R}_t &= \mathbb{1}\{|h_{t-1}(x'_t)| \leq M_t\}R_t & \overline{\mathbf{R}} &= \mathbb{E}[\overline{R}_t \,|\, \Xi_{t-1}] \\
\underline{R}_t &= \mathbb{1}\{|h_{t-1}(x'_t)| \geq M_t\}R_t & \underline{\mathbf{R}} &= \mathbb{E}[\underline{R}_t \,|\, \Xi_{t-1}]
\end{aligned} \tag{68}
$$

Notice that:

$$
L_t = \underline{L}_t + \overline{L}_t \qquad R_t = \underline{R}_t + \overline{R}_t. \tag{69}
$$

For $t \in \mathbb{N}$, define the following variables:

$$
\begin{aligned}
\overline{h}_t &:= \left[I_{\mathbb{H}} - \eta_t\left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t + \lambda_t I_{\mathbb{H}}\right)\right] h_{t-1} + \eta_t\left(K(x'_t,\cdot) - LR_t r^{\alpha}\right) + \eta_t\left(\mathbf{LR} - LR_t\right) g_{t-1} \\
&= h_t + \eta_t\left[(1-\alpha)\underline{L}_t + \alpha\underline{R}_t\right] h_{t-1} \qquad (\text{Eq. 64 and Eq. 69}) \\
k_t &:= \overline{h}_t - (1-\eta_t\lambda_t)h_{t-1} \\
&= h_t - \left[I_{\mathbb{H}} - \eta_t\left((1-\alpha)\underline{L}_t + \alpha\underline{R}_t + \lambda_t I_{\mathbb{H}}\right)\right] h_{t-1} \\
&= \eta_t\left[-\left[(1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right] h_{t-1} + K(x'_t,\cdot) + \mathbf{LR}(g_{t-1}) - LR_t\left(r^{\alpha} + g_{t-1}\right)\right] \qquad (\text{Eq. 64 and Eq. 69}).
\end{aligned} \tag{70}
$$

Lemma 4. *Assume $t_0^{\theta} \geq 2a\left(b + 2C^2\right)$. For all $t \in \mathbb{N}$, $M_t \in \mathbb{R}_+ \cup \{+\infty\}$, we have:*

$$
\mathbb{E}\left[\left\|\overline{h}_t\right\|_{\mathbb{H}}^2 \,|\, \Xi_{t-1}\right] \leq (1-\eta_t\lambda_t)^2 \left\|h_{t-1}\right\|_{\mathbb{H}}^2 + 2C^2\eta_t^2\left(\frac{2+\alpha^2}{\alpha^2}\right). \tag{71}
$$

In particular, assume that $\frac{1}{2} \leq \theta \leq 1$ and $t_0^{\theta} \geq \max\{2ab, 2\gamma, \gamma + \frac{2\theta-1}{\gamma}\}$ where $\gamma = ab - (\theta - \frac{1}{2}) > 0$, and fix $B_1 = aC\sqrt{2\left(\frac{2+\alpha^2}{\alpha^2\gamma}\right)}$. Then $\|h_{t-1}\|_{\mathbb{H}} \geq B_1\bar{t}^{\frac{1}{2}-\theta}$ implies:

$$
\bar{t}^{\theta-\frac{1}{2}}\mathbb{E}\left[\left\|\overline{h}_t\right\|_{\mathbb{H}} \,|\, \Xi_{t-1}\right] \leq (\bar{t}-1)^{\theta-\frac{1}{2}} \left\|h_{t-1}\right\|_{\mathbb{H}}. \tag{72}
$$

Proof. For all $t \in \mathbb{N}$, let us define the random variable:

$$
\zeta_t := \left[(1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}} - \left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right)\right] h_{t-1} + (\mathbf{LR} - LR_t)g_{t-1} + (K(x'_t,\cdot) - LR_t r^{\alpha}).
$$

Given the definition of $\overline{h}_t$ in Eq. 70:

$$
\overline{h}_t = \left[I_{\mathbb{H}} - \eta_t((1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}} + \lambda_t I_{\mathbb{H}})\right] h_{t-1} + \eta_t\zeta_t. \tag{73}
$$

The independence of the incoming observations (x_t, x'_t) and the fact h_{t-1}, g_{t-1} are Ξ_{t-1}-measurable lead to:

$$
\begin{aligned}
\mathbb{E}[\zeta_t \,|\, \Xi_{t-1}] &= \mathbb{E}[K(x'_t,\cdot) - LR_t r^{\alpha}] = \mathbb{E}[K(x'_t,\cdot) - ((1-\alpha)r^{\alpha}(x_t)K(x_t,\cdot) + \alpha r^{\alpha}(x'_t)K(x'_t,\cdot)) \,|\, \Xi_{t-1}] \\
&= \mathbb{E}_{q(x')}[K(x',\cdot)] - \mathbb{E}_{p^{\alpha}(y)}[r^{\alpha}(y)K(y,\cdot)] = 0.
\end{aligned} \tag{74}
$$

Eq. 73 and the last observation implies:

$$
\mathbb{E}[\left\|\overline{h}_t\right\|_{\mathbb{H}}^2 \,|\, \Xi_{t-1}] = \left\|\left[I_{\mathbb{H}} - \eta_t((1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}} + \lambda_t I_{\mathbb{H}})\right] h_{t-1}\right\|_{\mathbb{H}}^2 + \eta_t^2\mathbb{E}[\|\zeta_t\|_{\mathbb{H}}^2 \,|\, \Xi_{t-1}]. \tag{75}
$$

The next step is to upperbound each of the terms in Eq. 75. For the first element of the sum we have:

$$
\begin{aligned}
&\left\|\left[I_{\mathbb{H}} - \eta_t((1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}} + \lambda_t I_{\mathbb{H}})\right] h_{t-1}\right\|_{\mathbb{H}}^2 \\
&= (1-\lambda_t\eta_t)^2 \left\|h_{t-1}\right\|_{\mathbb{H}}^2 - 2\eta_t(1-\eta_t\lambda_t)\langle\left[(1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}}\right] h_{t-1}, h_{t-1}\rangle_{\mathbb{H}} + \eta_t^2\left\|\left[(1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}}\right] h_{t-1}\right\|_{\mathbb{H}}^2 \\
&= (1-\lambda_t\eta_t)^2 \left\|h_{t-1}\right\|_{\mathbb{H}}^2 - 2\eta_t(1-\eta_t\lambda_t)\mathbb{E}\left[(1-\alpha)h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}^2(x'_t)\mathbb{1}\{|h_{t-1}(x'_t)| \leq M_t\} \,|\, \Xi_{t-1}\right] \\
&+ \eta_t^2\left\|\left[(1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}}\right] h_{t-1}\right\|_{\mathbb{H}}^2 \qquad (\text{Eq. 68 and the first point of Eq. 2}).
\end{aligned} \tag{76}
$$

We can upper bound the last term in the previous expression by:

$$\begin{aligned}
&\eta_t^2 \left\| \left[(1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}}\right] h_{t-1} \right\|_{\mathbb{H}}^2 \\
&\leq \eta_t^2 \mathbb{E}\left[\|(1-\alpha)h_{t-1}(x_t)K(x_t,\cdot)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}(x_t')K(x_t',\cdot)\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\}\|_{\mathbb{H}}^2 \mid \Xi_{t-1} \right] \\
&\quad \big(\text{Jensen's inequality}\big) \\
&\leq \eta_t^2 \mathbb{E}\Big[(1-\alpha)^2 K(x_t,x_t) h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} \mid \Xi_{t-1}\Big] \\
&+ 2\eta_t^2(1-\alpha)\alpha \mathbb{E}\Big[K(x_t,x_t')\,|h_{t-1}(x_t)h_{t-1}(x_t')|\,\mathbb{1}\{|h_{t-1}(x_t)|\,,|h_{t-1}(x_t')| \leq M_t\}\Big] \\
&+ \eta_t^2\alpha^2 \mathbb{E}\Big[K(x_t',x_t')h_{t-1}^2(x_t')\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\} \mid \Xi_{t-1}\Big] \\
&\leq \eta_t^2 C^2 \mathbb{E}\Big[(1-\alpha)^2 h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} \mid \Xi_{t-1}\Big] \\
&+ 2\eta_t^2(1-\alpha)\alpha \mathbb{E}\Big[\,|h_{t-1}(x_t)h_{t-1}(x_t')|\,\mathbb{1}\{|h_{t-1}(x_t)|\,,|h_{t-1}(x_t')| \leq M_t\} \mid \Xi_{t-1}\Big] \\
&+ \eta_t^2\alpha^2 \mathbb{E}\Big[h_{t-1}^2(x_t')\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\} \mid \Xi_{t-1}\Big] \quad \big(\text{Assumption 2}\big) \\
&= \eta_t^2 C^2 \mathbb{E}\Big[\,\big((1-\alpha)h_{t-1}(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}(x_t')\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\}\big)^2 \mid \Xi_{t-1}\Big] \\
&\leq \eta_t^2 C^2 \mathbb{E}\Big[(1-\alpha)h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}^2(x_t')\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\} \mid \Xi_{t-1}\Big] \\
&\quad \big(\text{Jensen's inequality given } 0 \leq \alpha \leq 1\big)
\end{aligned} \tag{77}$$

Let us continue with the second term of Eq. 75. We start with an upperbound for the following:

$$\begin{aligned}
&\mathbb{E}\Big[\,\|LR_t g_{t-1}\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\Big] \\
&= \mathbb{E}\Big[\langle (1-\alpha)K(x_t,\cdot)g_{t-1}(x_t) + \alpha K(x_t',\cdot)g_{t-1}(x_t'), (1-\alpha)K(x_t,\cdot)g_{t-1}(x_t) + \alpha K(x_t',\cdot)g_{t-1}(x_t')\rangle_{\mathbb{H}} \mid \Xi_{t-1}\Big] \quad \big(\text{Eq. 2}\big) \\
&= \mathbb{E}\Big[(1-\alpha)^2 K(x_t,x_t)g_{t-1}^2(x_t) + 2(1-\alpha)\alpha K(x_t,x_t')g_{t-1}(x_t)g_{t-1}(x_t') \\
&+ \alpha^2 K(x_t',x_t')g_{t-1}^2(x_t') \mid \Xi_{t-1}\Big] \quad \big(\text{The first point of Eq. 2}\big) \\
&\leq C^2 \mathbb{E}\Big[\,[(1-\alpha)g_{t-1}(x_t) + \alpha g_{t-1}(x_t')]^2 \mid \Xi_{t-1}\Big] \quad \big(\text{Assumption 2}\big) \\
&\leq C^2 \mathbb{E}\Big[(1-\alpha)g_{t-1}^2(x_t) + \alpha g_{t-1}^2(x_t') \mid \Xi_{t-1}\Big] \quad \big(\text{Jensen's inequality given } 0 \leq \alpha \leq 1\big) \\
&\leq C^2 \mathbb{E}_{p^\alpha(y)}[g_{t-1}^2(y)] \leq \frac{C^2}{\alpha^2} \quad \big(\text{Lemma 2}\big).
\end{aligned} \tag{78}$$

Then, the second term of Eq. 75 satisfies the inequality:

$$\begin{aligned}
&\mathbb{E}\left[\|\zeta_t\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\right] \\
&\leq 2\mathbb{E}\left[\left\|\left[\big((1-\alpha)\overline{\mathbf{L}} + \alpha\overline{\mathbf{R}}\big) - \big((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\big)\right] h_{t-1} + (\mathbf{LR} - LR_t)g_{t-1})\right\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\right] \\
&+ 2\mathbb{E}\left[\|K(x_t',\cdot) - LR_t r^\alpha\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\right] \quad \big(2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|_{\mathbb{H}}^2 + \|b\|_{\mathbb{H}}^2\big) \\
&\leq 2\mathbb{E}\Big[\,\left\|\big((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\big)\, h_{t-1} + LR_t g_{t-1}\right\|_{\mathbb{H}}^2 + \|K(x_t',\cdot)\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\Big] \\
&\quad \big(\text{After developing the norms and taking conditional expectations}\big) \\
&\leq 2\Big[2\mathbb{E}\Big[\,\left\|\big((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\big)\, h_{t-1}\right\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\Big] + 2\mathbb{E}\Big[\,\|LR_t g_{t-1}\|_{\mathbb{H}}^2 \mid \Xi_{t-1}\Big] + C^2\Big] \\
&\quad \big(2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|_{\mathbb{H}}^2 + \|b\|_{\mathbb{H}}^2 \text{ and Assumption 2}\big) \\
&\leq 2C^2\left[2\mathbb{E}\Big[(1-\alpha)h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}^2(x_t')\mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\} \mid \Xi_{t-1}\Big] + \frac{2}{\alpha^2} + 1\right] \quad \big(\text{Eq. 78}\big).
\end{aligned} \tag{79}$$

By putting together Expressions 75-79:

$$\begin{aligned}
&\mathbb{E}[\|\bar{h}_t\|_{\mathbb{H}}^2 \mid \Xi_{t-1}] \\
&\leq (1-\lambda_t\eta_t)^2 \|h_{t-1}\|_{\mathbb{H}}^2 \\
&- \eta_t\left(2-2\eta_t\lambda_t - 5C^2\eta_t\right)\mathbb{E}\left[(1-\alpha)h_{t-1}^2(x_t)\mathbb{1}\{|h_{t-1}(x_t)|\leq M_t\} + \alpha h_{t-1}^2(x_t')\mathbb{1}\{|h_{t-1}(x_t')|\leq M_t\} \mid \Xi_{t-1}\right] \\
&+ 2C^2\eta_t^2\left(\frac{2+\alpha^2}{\alpha^2}\right) \\
&\leq (1-\lambda_t\eta_t)^2\|h_{t-1}\|_{\mathbb{H}}^2 + 2C^2\eta_t^2\left(\frac{2+\alpha^2}{\alpha^2}\right).
\end{aligned} \tag{80}$$

The last is a consequence of the hypothesis $t_0^\theta \geq 2a(b+2C^2)$, which implies $2-2\eta_t\lambda_t - 5C^2\eta_t \geq 0$ for all $t \in \mathbb{N}$. Therefore, we obtain the first inequality of Lemma 4.

The second point of that lemma depends on the following inequality:

$$(1-\frac{1}{\bar{t}})^{1-2\theta}(1-\frac{ab}{\bar{t}})^2 \leq (1-\frac{\gamma}{\bar{t}}), \tag{81}$$

where $\theta \in [\frac{1}{2}, 1]$ and $\bar{t} = t + t_0$ and $t_0 \geq \max\{2ab, 2\gamma, \gamma + \frac{2\theta-1}{\gamma}\}$, where $\gamma = ab - \frac{(2\theta-1)}{2}$.

Inequality 81 can be verified as follows:

$$\begin{aligned}
\log\left[(1-\frac{1}{\bar{t}})^{1-2\theta}(1-\frac{ab}{\bar{t}})^2(1-\frac{\gamma}{\bar{t}})^{-1}\right] &\leq -(2\theta-1)\log\left(1-\frac{1}{\bar{t}}\right) + 2\log\left(1-\frac{ab}{\bar{t}}\right) - \log\left(1-\frac{\gamma}{\bar{t}}\right) \\
&\leq \frac{(2\theta-1)}{\bar{t}} + \frac{(2\theta-1)}{\bar{t}^2} - \frac{2ab}{\bar{t}} + \frac{\gamma}{\bar{t}} + \frac{\gamma^2}{\bar{t}^2} \\
&= -\frac{\gamma}{\bar{t}} + \frac{2\theta-1+\gamma^2}{\bar{t}^2} = \frac{\gamma}{t}\left(\frac{\gamma}{\bar{t}}-1\right) + \frac{2\theta-1}{t^2} \leq 0,
\end{aligned}$$

where for the second equality we have used the inequalities $\log(1-x) \leq -x$ for all $x \in [0,1]$ and $\log(1-x) \geq -x-x^2$ for $x \in [0, \frac{1}{2}]$.

By using Inequalities 80 and 81, we can verify:

$$\begin{aligned}
\mathbb{E}\left[\left(1-\frac{1}{\bar{t}}\right)^{1-2\theta}\|\bar{h}_t\|_{\mathbb{H}}^2 - \|h_{t-1}\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] &\leq \left(1-\frac{1}{\bar{t}}\right)^{1-2\theta}(1-\frac{ab}{\bar{t}})^2\|h_{t-1}\|_{\mathbb{H}}^2 - \|h_{t-1}\|_{\mathbb{H}}^2 + \frac{2a^2C^2}{\bar{t}^{2\theta}}\left(\frac{2+\alpha^2}{\alpha^2}\right) \\
&\leq -\frac{\gamma}{\bar{t}}\|h_{t-1}\|_{\mathbb{H}}^2 + \frac{2a^2C^2}{\bar{t}^{2\theta}}\left(\frac{2+\alpha^2}{\alpha^2}\right)(\text{Thanks to the hypothesis on } t_0 \text{ and } \theta).
\end{aligned} \tag{82}$$

If $\|h_{t-1}\|_{\mathbb{H}} \geq B_1\bar{t}^{\frac{1}{2}-\theta}$, we can verify:

$$\mathbb{E}\left[\left(1-\frac{1}{\bar{t}}\right)^{1-2\theta}\|\bar{h}_t\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] \leq \left(1-\frac{\gamma}{\bar{t}}\right)\|h_{t-1}\|_{\mathbb{H}}^2 + \frac{2a^2C^2}{\bar{t}^{2\theta}}\left(\frac{2+\alpha^2}{\alpha^2}\right) \leq \left(1-\frac{\gamma}{\bar{t}}\right)\|h_{t-1}\|_{\mathbb{H}}^2 + \frac{B_1^2\gamma}{\bar{t}^{2\theta}} \leq \|h_{t-1}\|_{\mathbb{H}}^2. \tag{83}$$

After applying Jensen's inequality, we obtain the desired result. □

Lemma 5. *Assume $t_0^\theta \geq a(C^2+b)$ and $t_0^{1-\theta} \geq b(\alpha(M_t+1)+3)$; then,*

$$\|k_t\|_{\mathbb{H}} \leq \frac{Cab^{-1}}{\alpha\bar{t}^{2\theta-1}} \quad \textit{and} \quad \mathbb{E}\left[\|k_t\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] \leq \frac{3\eta_t^2C^2}{\alpha^2}\left[\alpha^2(M_t^2+1)+1\right]. \tag{84}$$

Proof. Let us start with the following inequality:

$$\begin{aligned}
&\left\| \left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right) h_{t-1} \right\|_{\mathbb{H}}^2 \\
&\leq (1-\alpha)^2 K(x_t,x_t) h_{t-1}^2(x_t) \mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} \\
&+ 2\alpha(1-\alpha) K(x_t,x_t') \left|h_{t-1}(x_t)\right| \left|h_{t-1}(x_t')\right| \mathbb{1}\{\left|h_{t-1}(x_t')\right|, \left|h_{t-1}(x_t')\right| \leq M_t\} \\
&+ \alpha^2 K(x_t',x_t') h_{t-1}^2(x_t') \mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\} \quad \left(\text{First point of 2 and after developing the norm}\right) \\
&\leq C^2 \left((1-\alpha) h_{t-1}(x_t) \mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}(x_t') \mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\}\right)^2 \quad \left(\text{Assumption 2}\right) \\
&\leq C^2 \left[(1-\alpha) h_{t-1}^2(x_t) \mathbb{1}\{|h_{t-1}(x_t)| \leq M_t\} + \alpha h_{t-1}^2(x_t') \mathbb{1}\{|h_{t-1}(x_t')| \leq M_t\}\right] \quad \left(\text{Jensen's inequality given } 0 \leq \alpha \leq 1\right) \\
&\leq C^2 M_t^2.
\end{aligned} \tag{85}$$

Moreover, by exploiting the hypothesis $r^\alpha \in \mathbb{H}$ and after following the same line of argumentation as in the previous inequality, we verify:

$$\begin{aligned}
\|LR_t(r^\alpha + g_{t-1})\|_{\mathbb{H}}^2 &\leq C^2 \left((1-\alpha)(r^\alpha + g_{t-1})^2(x_t) + \alpha (r^\alpha + g_{t-1})^2(x_t')\right) \\
&\leq C^2 (1-\alpha) \mathbb{E}_{p(x)}[(r^\alpha + g_{t-1})^2(x)] + \alpha \mathbb{E}_{q(x')}[(r^\alpha + g_{t-1})^2(x')] \\
&= C^2 \left\|r^\alpha + g_{t-1}\right\|_{L^2_{p^\alpha}}^2,
\end{aligned} \tag{86}$$

which implies:

$$\|LR_t(r^\alpha + g_{t-1})\|_{\mathbb{H}} \leq C \left\|r^\alpha + g_{t-1}\right\|_{L^2_{p^\alpha}} \leq C \left(\|r^\alpha\|_{L^2_{p^\alpha}} + \|g_{t-1}\|_{L^2_{p^\alpha}}\right) \leq \frac{2C}{\alpha}. \tag{87}$$

The last line is a consequence of the first point of Lemma 2 and the fact $r^\alpha \leq \frac{1}{\alpha}$.

We can now upperbound the norm $\|k_t\|_{\mathbb{H}}$:

$$\begin{aligned}
\|k_t\|_{\mathbb{H}} &= \eta_t \left[\left\|\left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right) h_{t-1}\right\|_{\mathbb{H}} + \|K(x_t,\cdot)\|_{\mathbb{H}} + \|\mathbf{LR} g_{t-1}\|_{\mathbb{H}} + \|LR_t\left(g_{t-1} + r^\alpha\right)\|_{\mathbb{H}}\right] \\
&= \eta_t \left[C(M_t + 1 + \frac{1}{\alpha} + \frac{2}{\alpha})\right] \quad \left(\text{Eq. 85, Assumption 2, Lemma 3, Eq. 87}\right) \\
&\leq \frac{C\eta_t}{\alpha}\left[\alpha(M_t+1)+3\right] \\
&\leq \frac{\eta_t C}{\alpha\lambda_t} \quad \left(\text{Hypothesis } t_0^{1-\theta} \geq b\left[\alpha\left(M_t+1\right)+3\right] \text{ implies } \alpha(M_t+1)+3 \leq \frac{1}{\lambda_t}\right) \\
&\leq \frac{Cab^{-1}}{\alpha \bar{t}^{2\theta-1}}.
\end{aligned}$$

On the other hand, we obtain for the expected norm:

$$\begin{aligned}
\mathbb{E}\left[\|k_t\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] &\leq 3\eta_t^2 \Bigg[\mathbb{E}\left[\left\|\left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right) h_{t-1}\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] + \mathbb{E}\left[\left\|K(x_t',\cdot) - LR_t\left(r^\alpha\right)\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] \\
&\quad + \mathbb{E}\left[\left\|(\mathbf{LR} - LR_t)\, g_{t-1}\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right]\Bigg] \quad \left(2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|_{\mathbb{H}}^2 + \|b\|_{\mathbb{H}}^2\right) \\
&\leq 3\eta_t^2 \Bigg[\mathbb{E}\left[\left\|\left((1-\alpha)\overline{L}_t + \alpha\overline{R}_t\right) h_{t-1}\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] + \mathbb{E}\left[\left\|K(x_t',\cdot)\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right] \\
&\quad + \mathbb{E}\left[\left\|LR_t g_{t-1}\right\|_{\mathbb{H}}^2 \middle| \Xi_{t-1}\right]\Bigg] \quad \left(\text{After developing the norms and taking conditional expectations}\right) \\
&\leq 3\eta_t^2 C^2 \left[M_t^2 + 1 + \|g_{t-1}\|_{L^2_{p^\alpha}}^2\right] \quad \left(\text{Eq. 85 and the same line of reasoning that in Eq. 86}\right) \\
&\leq \frac{3\eta_t^2 C^2}{\alpha^2}\left[\alpha^2(M_t^2+1)+1\right]
\end{aligned}$$

□

Lemma 6. *For all $t \in \mathbb{N}$, assume $M_t \geq \frac{2C^2 ab^{-1}\bar{t}^{1-2\theta}}{\alpha}$, $t_0^\theta \geq 2a(C^2+b)$ and $t_0^{1-\theta} \geq b\left(\alpha(M_t+1)+3\right)$, then*

$$\|h_t\|_{\mathbb{H}} \leq \left\|\overline{h}_t\right\|_{\mathbb{H}}. \tag{88}$$

Proof. We start with the following inequality that relates $\bar{h}_{t-1}$ and k_t, take $x \in \mathcal{X}$ such that $h_{t-1}(x) \geq M_t$, then:

$$\begin{aligned}\bar{h}_t(x) &= (1-\lambda_t\eta_t)h_{t-1}(x) + k_t(x) \quad (\text{Eq. } 70) \\ &\geq (1-\lambda_t\eta_t)h_{t-1}(x) - \frac{C^2ab^{-1}}{\alpha\bar{t}^{2\theta-1}} \quad (\text{Lemma } 5) \\ &\geq (1-\lambda_t\eta_t)h_{t-1}(x) - \frac{1}{2}h_{t-1}(x) \quad \left(\text{As a consequence of } M_t \geq \frac{2C^2ab^{-1}\bar{t}^{1-2\theta}}{\alpha}\right) \\ &\geq C^2\eta_t\frac{h_{t-1}(x)}{2}.\end{aligned} \tag{89}$$

The last identity is a consequence of assumption $t_0^\theta \geq 2a(C^2+b)$ which implies $1-\eta_t\lambda_t - \frac{C^2\eta_t}{2} \geq \frac{1}{2}$.
Suppose we have $h_{t-1}(x_t) \geq M_t$ and $h_{t-1}(x'_t) \geq M_t$, then we have:

$$h_t = \bar{h}_t - \eta_t\left[(1-\alpha)L_t + \alpha R_t\right]h_{t-1} \quad (\text{Eq. } 70) \tag{90}$$

$$\begin{aligned}\|h_t\|^2_{\mathbb{H}} = \langle h_t, h_t\rangle_{\mathbb{H}} &= \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - 2\eta_t\langle \bar{h}_t, \left[(1-\alpha)L_t + \alpha R_t\right]h_{t-1}\rangle_{\mathbb{H}} + \eta_t^2\left\|\left[(1-\alpha)L_t+\alpha R_t\right]h_{t-1}\right\|^2_{\mathbb{H}} \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - 2\eta_t\Big((1-\alpha)\bar{h}_t(x_t)h_{t-1}(x_t) + \alpha\bar{h}_t(x'_t)h_{t-1}(x'_t)\Big) + \eta_t^2\Big[(1-\alpha)^2\left\|L_th_{t-1}\right\|^2_{\mathbb{H}} \\ &+\alpha^2\left\|R_th_{t-1}\right\|^2_{\mathbb{H}} + 2(1-\alpha)\alpha K(x_t,x'_t)h_{t-1}(x_t)h_{t-1}(x'_t)\Big] \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - \eta_t^2C^2\left[(1-\alpha)h^2_{t-1}(x_t) + \alpha h^2_{t-1}(x'_t)\right] + 2\eta_t^2C^2\alpha(1-\alpha)\left|h_{t-1}(x_t)h_{t-1}(x'_t)\right| \\ &+\eta_t^2\left[(1-\alpha)^2\left\|L_th_{t-1}\right\|^2_{\mathbb{H}} + \alpha^2\left\|R_th_{t-1}\right\|^2_{\mathbb{H}}\right] \quad (\text{Eq. } 90 \text{ and Eq. } 89) \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - \eta_t^2C^2\left[(1-\alpha)h_{t-1}(x_t) + \alpha h_{t-1}(x'_t)\right]^2 + 2\eta_t^2C^2\alpha(1-\alpha)\left|h_{t-1}(x_t)h_{t-1}(x'_t)\right| \\ &+\eta_t^2\left[(1-\alpha)^2\left\|L_th_{t-1}\right\|^2_{\mathbb{H}} + \alpha^2\left\|R_th_{t-1}\right\|^2_{\mathbb{H}}\right] \quad (\text{Jensen's inequality given } 0\leq\alpha\leq 1) \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - \eta_t^2C^2\left[(1-\alpha)^2h^2_{t-1}(x_t) + \alpha^2h^2_{t-1}(x'_t)\right] + \eta_t^2\left[(1-\alpha)^2\left\|L_th_{t-1}\right\|^2_{\mathbb{H}} + \alpha^2\left\|R_th_{t-1}\right\|^2_{\mathbb{H}}\right] \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} + \eta_t^2(1-\alpha)^2h^2_{t-1}(x_t)\left[K(x_t,x_t)-C^2\right] + \eta_t^2\alpha^2h^2_{t-1}(x'_t)\left[K(x'_t,x'_t)-C^2\right] \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} \quad (\text{Assumption } 2).\end{aligned} \tag{91}$$

Let us continue with the case $h_{t-1}(x_t) \geq M_t$ and $h_{t-1}(x'_t) < M_t$, then we have:

$$h_t = \bar{h}_t - (1-\alpha)\eta_tL_th_{t-1} \quad (\text{Eq. } 70).$$

Then by following the same line of argumentation than in the previous point, we get:

$$\begin{aligned}\|h_t\|^2_{\mathbb{H}} &= \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - 2(1-\alpha)\eta_t\langle L_th_{t-1}, \bar{h}_t\rangle_{\mathbb{H}} + \eta_t^2(1-\alpha)^2\left\|L_th_{t-1}\right\|^2_{\mathbb{H}} \\ &= \left\|\bar{h}_t\right\|^2_{\mathbb{H}} - 2(1-\alpha)\eta_th_{t-1}(x_t)\bar{h}_t(x_t) + \eta_t^2(1-\alpha)^2h^2_{t-1}(x_t)K(x_t,x_t) \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} + \eta_t^2(1-\alpha)^2h^2_{t-1}(x_t)K(x_t,x_t) - (1-\alpha)C^2\eta_t^2h^2_{t-1}(x_t) \quad \left(\bar{h}_t(x_t) \geq C^2\eta_t\frac{h_{t-1}(x_t)}{2}\right) \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} + \eta_t^2(1-\alpha)^2h^2_{t-1}(x_t)K(x_t,x_t) - (1-\alpha)^2C^2\eta_t^2h^2_{t-1}(x_t) \\ &= \left\|\bar{h}_t\right\|^2_{\mathbb{H}} + (1-\alpha)^2\eta_t^2h^2_{t-1}(x_t)\left(K(x_t,x_t)-C^2\right) \\ &\leq \left\|\bar{h}_t\right\|^2_{\mathbb{H}} \quad (\text{Assumption } 2).\end{aligned}$$

The case $h_{t-1}(x_t) < M_t$ and $h_{t-1}(x'_t) \geq M_t$ can be solve in a symmetric way. Finally, for $h_{t-1}(x_t) < M_t$ and $h_{t-1}(x'_t) < M_t$, the inequality follows directly. ☐

Lemma 7. *Assume $\theta \in [\frac{1}{2},1]$,$t_0 \geq 3$, $b = a^{-1}$, $t_0^\theta \geq 2 + 4C^2a$ and $t_0^{1-\theta} \geq 8b$, Then, with probability at least $1-\delta$:*

$$\sup_{0\leq k\leq t}\|h_k\|_{\mathbb{H}}(k+t_0+1)^{\theta-\frac{1}{2}} \leq \frac{aC}{\alpha}\left[5at_0^{\frac{1}{2}-\theta} + (14Ca^2+18)\sqrt{\log(\bar{t})}\right]\log\left(\frac{2}{\delta}\right) := B_{t,\delta}. \tag{92}$$

Proof. Let us start by verifying that the hypothesis of Lemmas 4, 5, and 6.

First $t_0^{\theta} \geq 2+4C^2a = 2a(b+2C^2)$ and $\gamma = ab - \left(\theta - \frac{1}{2}\right) \in [\frac{1}{2}, 1]$, where we we have used the assumption $ab = 1$. The assumption $t_0 \geq 3$ implies $t_0 \geq \max\left(2ab, 2\gamma, \gamma + \frac{2\theta-1}{\gamma}\right)$.

Finally, if we fix $M_t = \frac{2C^2ab^{-1}\bar{t}^{1-2\theta}}{\alpha}$, the fact that $t_0^{\theta} \geq 4C^2a$ and $t_0^{1-\theta} \geq 8b$ implies:

$$t_0^{1-\theta} \geq \frac{t_0^{1-\theta}}{2} + \frac{t_0^{1-\theta}}{2} = \frac{t_0^{1-\theta}}{2} + \frac{t_0^{1-2\theta}(t_0^{\theta})}{2} \geq b\left(4 + 2C^2ab^{-1}t_0^{1-2\theta}\right) \geq b\left(4+\alpha M_t\right) \geq b\left(3+\alpha(M_t+1)\right).$$

Then, the required assumptions are satisfied.

Take $i \in \mathbb{N}$, if $\|h_{i-1}\|_{\mathbb{H}} \geq B_1\bar{t}^{\frac{1}{2}-\theta}$, where $B_1 = aC\sqrt{2\left(\frac{2+\alpha^2}{\alpha^2\gamma}\right)}$, then:

$$\begin{aligned} \|h_i\|_{\mathbb{H}} &\leq \left\|\bar{h}_i\right\|_{\mathbb{H}} \quad (\text{Lemma 6}) \\ &= \left\|\bar{h}_i\right\|_{\mathbb{H}} - \mathbb{E}\left[\left\|\bar{h}_i\right\|_{\mathbb{H}} \mid \Xi_{i-1}\right] + \mathbb{E}\left[\left\|\bar{h}_i\right\|_{\mathbb{H}} \mid \Xi_{i-1}\right] \\ &\leq \xi_i + \left[1 - \frac{1}{i+t_0}\right]^{\theta-\frac{1}{2}} \|h_{i-1}\|_{\mathbb{H}} \quad (\text{Second part of Lemma 4}), \end{aligned} \tag{93}$$

where $\xi_i := \left\|\bar{h}_i\right\|_{\mathbb{H}} - \mathbb{E}\left[\left\|\bar{h}_i\right\|_{\mathbb{H}} \mid \Xi_{i-1}\right]$.

Notice that the stochastic process $\{\xi_k\}_{k\in\mathbb{N}}$ defines a martingale difference sequence, which additionally satisfies the following inequalities:

$$\begin{aligned} |\xi_i| &\leq \left\|\bar{h}_i - \mathbb{E}\left[\bar{h}_i \mid \Xi_{i-1}\right]\right\|_{\mathbb{H}} \\ &= \|k_i - \mathbb{E}\left[k_i \mid \Xi_{i-1}\right]\|_{\mathbb{H}} \quad (\text{Eq. 70}) \\ &\leq \|k_i\|_{\mathbb{H}} + \|\mathbb{E}\left[k_i \mid \Xi_{i-1}\right]\|_{\mathbb{H}} \\ &\leq \|k_i\|_{\mathbb{H}} + \mathbb{E}\left[\|k_i\|_{\mathbb{H}} \mid \Xi_{i-1}\right] \quad (\text{Jensen's inequality}) \\ &\leq \frac{2Cab^{-1}}{\alpha(i+t_0)^{2\theta-1}} \quad (\text{Lemma 5}). \end{aligned} \tag{94}$$

In a similar manner, we can verify:

$$\begin{aligned} \mathbb{E}\left[\xi_i^2 \mid \Xi_{i-1}\right] &= \mathbb{E}\left[\|k_i - \mathbb{E}\left[k_i \mid \Xi_{i-1}\right]\|_{\mathbb{H}}^2 \mid \Xi_{i-1}\right] \\ &\leq \mathbb{E}\left[\|k_i\|_{\mathbb{H}}^2 \mid \Xi_{i-1}\right] - \|\mathbb{E}\left[k_i \mid \Xi_{i-1}\right]\|_{\mathbb{H}}^2 \\ &\leq \mathbb{E}\left[\|k_i\|_{\mathbb{H}}^2 \mid \Xi_{i-1}\right] \\ &\leq \frac{3\eta_i^2C^2}{\alpha^2}\left[\alpha^2(M_i^2+1)+1\right] \quad (\text{Lemma 5}) \\ &= \frac{3\eta_i^2C^2}{\alpha^2}\left(\alpha^2\left(\frac{4C^4a^2b^{-2}(t+i)^{2(1-2\theta)}}{\alpha^2}+1\right)+1\right) \\ &\leq \frac{3\eta_i^2C^2}{\alpha^2}(4C^2a^2b^{-2}+2) \\ &= \frac{12\eta_i^2C^2}{\alpha^2}(C^2a^2b^{-2}+\frac{1}{2}) \\ &\leq \frac{12\eta_i^2C^2}{\alpha^2}(Cab^{-1}+1)^2 \end{aligned} \tag{95}$$

Notice that the sequence $\left\{(k+t_0)^{\theta-\frac{1}{2}}\xi_k\right\}$ defines a difference martingale as well, which satisfies the inequalities:

$$\begin{aligned} \left|(i+t_0)^{\theta-\frac{1}{2}}\xi_i\right| &\leq \frac{2Ca^2(i+t_0)^{\frac{1}{2}-\theta}}{\alpha} \quad (\text{Eq. 94}) \\ &\leq \frac{2Ca^2t_0^{\frac{1}{2}-\theta}}{\alpha} \end{aligned} \tag{96}$$

$$\begin{aligned}\sum_{k=1}^{t}\mathbb{E}\left[\left((k+t_0)^{\theta-\frac{1}{2}}\xi_k\right)^2 \mid \Xi_{k-1}\right] &\leq \frac{12a^2C^2}{\alpha^2}(Cab^{-1}+1)^2\sum_{k=1}^{t}(k+t_0)^{-1} \quad (\text{Eq. } 95) \\ &\leq \frac{12a^2C^2}{\alpha^2}(Cab^{-1}+1)^2\log(1+\frac{t}{t_0}).\end{aligned} \tag{97}$$

Let us define the term:

$$\nu_i = \sum_{j=1}^{i}\xi_j(j+t_0)^{\theta-\frac{1}{2}}\mathbb{1}\{\|h_{j-1}\|_{\mathbb{H}} \geq B_1(j+t_0)^{\frac{1}{2}-\theta}\}. \tag{98}$$

Inequalities 96 and 97 imply the hypothesis of Preposition A.3 in Tarrès and Yao (2014) (Lemma 9) are satisfied. Then, the probability of the event Δ, $P(\Delta) \geq 1-\delta$, where:

$$\begin{aligned}\Delta = \Big\{ \sup_{1\leq i\leq t}|\nu_i| &\leq 2\left(\frac{2Ca^2t_0^{\frac{1}{2}-\theta}}{3\alpha} + \frac{2\sqrt{3}aC}{\alpha}(Cab^{-1}+1)\sqrt{\log(1+\frac{t}{t_0})}\right)\log\left(\frac{2}{\delta}\right) \\ &\leq \frac{4aC}{\alpha}\left(\frac{at_0^{\frac{1}{2}-\theta}}{3} + \sqrt{3}(Cab^{-1}+1)\sqrt{\log\left(1+\frac{t}{t_0}\right)}\right)\log\left(\frac{2}{\delta}\right) \\ &= \frac{4aC}{\alpha}\left(\frac{at_0^{\frac{1}{2}-\theta}}{3} + \sqrt{3}(Ca^2+1)\sqrt{\log\left(1+\frac{t}{t_0}\right)}\right)\log\left(\frac{2}{\delta}\right)\Big\}.\end{aligned} \tag{99}$$

Assume that the event Δ holds and let u_k for all $k \in \mathbb{N}$ be:

$$u_k = \|h_k\|(k+t_0)^{\theta-\frac{1}{2}}. \tag{100}$$

For all the elements $k \leq t$, let:

$$m = \max\{j \leq k : \|h_j\|_{\mathbb{H}} < B_1(j+t_0+1)^{\frac{1}{2}-\theta}\}. \tag{101}$$

If $m < k$, then:

$$\begin{aligned}u_{m+1} &\leq \left[\left(\frac{m+t_0}{m+1+t_0}\right)^{\theta-\frac{1}{2}}\|h_m\|_{\mathbb{H}} + |\xi_{m+1}|\right](m+1+t_0)^{\theta-\frac{1}{2}} \quad (\text{Eq. } 93) \\ &< \left[\left(\frac{m+t_0}{m+t_0+1}\right)^{\theta-\frac{1}{2}}(m+t_0+1)^{\frac{1}{2}-\theta}B_1 + \frac{2Ca^2}{\alpha}(m+1+t_0)^{1-2\theta}\right](m+1+t_0)^{\theta-\frac{1}{2}} \quad (\text{Eq. } 94 \text{ and Eq. } 101) \\ &\leq aC\sqrt{2\left(\frac{2+\alpha^2}{\alpha^2\gamma}\right)} + \frac{2Ca^2}{\alpha}t_0^{\frac{1}{2}-\theta} \leq aC\sqrt{4\left(\frac{2+\alpha^2}{\alpha^2}\right)} + \frac{2Ca^2}{\alpha}t_0^{\frac{1}{2}-\theta} \quad (\gamma \in [\tfrac{1}{2},1]) \\ &\leq \frac{aC}{\alpha}\left(\sqrt{4(2+\alpha^2)} + 2at_0^{\frac{1}{2}-\theta}\right) \leq \frac{2aC}{\alpha}\left(\sqrt{3}+at_0^{\frac{1}{2}-\theta}\right).\end{aligned} \tag{102}$$

Given Expr. 93, we can verify:

$$(i+t_0)^{\theta-\frac{1}{2}}\|h_i\|_{\mathbb{H}} \leq (i+t_0)^{\theta-\frac{1}{2}}\xi_i + (i-1+t_0)^{\theta-\frac{1}{2}}\|h_{i-1}\|_{\mathbb{H}} \tag{103}$$

Then by recursion and given by the definition of Expr. 98, we get:

$$u_k \leq u_{m+1} + \nu_k - \nu_{m+1}. \tag{104}$$

For δ sufficiently small, we have:

$$
\begin{aligned}
u_k &\leq u_{m+1} + |\nu_k| + |\nu_{m+1}| \\
&\leq \frac{2aC}{\alpha}\left[\left(\frac{4}{3}+1\right)at_0^{\frac{1}{2}-\theta} + \sqrt{3} + 4\sqrt{3}(Ca^2+1)\sqrt{\log\left(1+\frac{t}{t_0}\right)}\right]\log\left(\frac{2}{\delta}\right) \quad \text{(Eq. 99 and Eq. 102)} \\
&= \frac{aC}{\alpha}\left[\left(\frac{14}{3}\right)at_0^{\frac{1}{2}-\theta} + 2\sqrt{3} + 8\sqrt{3}(Ca^2+1)\sqrt{\log\left(1+\frac{t}{t_0}\right)}\right]\log\left(\frac{2}{\delta}\right) \quad \text{(Eq. 99)} \\
&\leq \frac{aC}{\alpha}\left[5at_0^{\frac{1}{2}-\theta} + 4 + 14(Ca^2+1)\sqrt{\log\left(\frac{\bar{t}}{t_0}\right)}\right]\log\left(\frac{2}{\delta}\right) \\
&\leq \frac{aC}{\alpha}\left[5at_0^{\frac{1}{2}-\theta} + (14Ca^2+18)\sqrt{\log\left(\bar{t}\right)}\right]\log\left(\frac{2}{\delta}\right),
\end{aligned}
\tag{105}
$$

where the fact that $t_0 \geq 3$ implies $\sqrt{\log(t+t_0)} \geq 1$, meaning $4 + 14(Ca^2+1)\sqrt{\log\left(\frac{\bar{t}}{t_0}\right)} \leq (14Ca^2 + 18)\sqrt{\log(\bar{t})}\log\left(\frac{2}{\delta}\right)$. □

Proof of Theorem 6.

Proof. Let us start with a basic inequality that will be useful during the proof; suppose f is Ξ_{j-1}-measurable, then we have:

$$
\begin{aligned}
\mathbb{E}\left[\|LR_j f\|_{\mathbb{H}}^2 \,|\, \Xi_{j-1}\right] &= \mathbb{E}\left[\langle(1-\alpha)f(x_j)K(x_j,\cdot) + \alpha f(x_j')K(x_j',\cdot), (1-\alpha)f(x_j)K(x_j,\cdot) + \alpha f(x_j')K(x_j',\cdot)\rangle_{\mathbb{H}} \,|\, \Xi_{j-1}\right] \\
&= \mathbb{E}\left[(1-\alpha)^2 K(x_j,x_j)f^2(x_j) + 2(1-\alpha)\alpha K(x_j,x_j')f(x_j)f(x_j') + \alpha^2 K(x_j',x_j')f^2(x_j') \,|\, \Xi_{j-1}\right] \\
&\leq C^2\mathbb{E}\left[\left((1-\alpha)f(x_j) + \alpha f(x_j')\right)^2 \,|\, \Xi_{j-1}\right] \quad \text{(Assumption 2)} \\
&\leq C^2\mathbb{E}\left[(1-\alpha)f^2(x_j) + \alpha f^2(x_j') \,|\, \Xi_{j-1}\right] \quad \text{(Jensen's inequality)} \\
&= C^2\|f\|_{L^2_{p^\alpha}}^2.
\end{aligned}
\tag{106}
$$

Notice that for $f \in \mathbb{H}$:

$$
\begin{aligned}
\|LR_j f\|_{\mathbb{H}} &\leq (1-\alpha)\left|\langle K(x_j,\cdot), f\rangle_{\mathbb{H}}\right| \|K(x_j,\cdot)\|_{\mathbb{H}} + \alpha\left|\langle K(x_j',\cdot), f\rangle_{\mathbb{H}}\right| \|K(x_j,\cdot)\|_{\mathbb{H}} \\
&\leq \left[(1-\alpha)\|K(x_j,\cdot)\|_{\mathbb{H}}^2 + \alpha\left\|K(x_j',\cdot)\right\|_{\mathbb{H}}^2\right]\|f\|_{\mathbb{H}} \quad \text{(Cauchy–Schwarz inequality)} \\
&\leq C^2\|f\|_{\mathbb{H}},
\end{aligned}
\tag{107}
$$

which implies $\|LR_j\| \leq C^2$.

Fix $t \in \mathbb{N}$, $\delta \in [0,1]$, and let

$$
B_{t,\delta} = \frac{aC}{\alpha}\left[5at_0^{\frac{1}{2}-\theta} + (14Ca^2+18)\sqrt{\log\left(\bar{t}\right)}\right]\log\left(\frac{2}{\delta}\right) \tag{108}
$$

and the following stochastic process:

$$
\Upsilon_j = \eta_j\bar{\Pi}_{j+1}^t\epsilon_j\mathbb{1}\{\|h_{j-1}\|_{\mathbb{H}}(j+t_0)^{\theta-\frac{1}{2}} \leq B_{t,\delta}\}. \tag{109}
$$

Verifying that the sequence $\{\Upsilon_j\}_{j\in\mathbb{N}}$ is a difference martingale is easy. The idea to finish the proof is to apply Lemma 9 to the sequence Υ_j, this means, we should show the existence of $M > 0$ and $\sigma^2 > 0$ such that $\|\Upsilon_j\|_{L^2_{p^\alpha}} \leq M$ and $\sum_{j=1}^t \mathbb{E}\left[\|\Upsilon_j\|_{L^2_{p^\alpha}}^2 \,|\, \Xi_{j-1}\right] \leq \sigma^2$.

Let us start by identifying σ^2. Suppose $\|h_{j-1}\|_{\mathbb{H}} \leq B_{t,\delta}(j+t_0)^{\frac{1}{2}-\theta}$, then by using the decomposition $f_j =$

$r^\alpha + g_j + h_j$ and the inequalities stated in Lemma 2 we have:

$$
\begin{aligned}
\mathbb{E}\left[\|\epsilon_j\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] &= \mathbb{E}\left[\left\|(\mathcal{L}_K - LR_j)f_{j-1} + K(x'_j,\cdot) - \mathcal{L}_K r^\alpha\right\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \\
&\leq \mathbb{E}\left[\left\|K(x'_j,\cdot) - LR_j f_{j-1}\right\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \quad \text{(After developing the norm and taking conditional expectations)} \\
&\leq \mathbb{E}\left[\left\|K(x'_j,\cdot) - LR_j r^\alpha - LR_j g_{j-1} - LR_j h_{j-1}\right\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \\
&\leq 4\mathbb{E}\left[\left\|K(x'_j,\cdot)\right\|_{\mathbb{H}}^2 + \|LR_j r^\alpha\|_{\mathbb{H}}^2 + \|LR_j g_{j-1}\|_{\mathbb{H}}^2 + \|LR_j h_{j-1}\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \quad \left(2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|_{\mathbb{H}}^2 + \|b\|_{\mathbb{H}}^2\right) \\
&\leq 4\left[C^2 + \frac{C^2}{\alpha^2} + \frac{C^2}{\alpha^2} + \mathbb{E}\left[\|LR_j\|^2 \|h_{j-1}\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right]\right] \quad \text{(Assumption 2 , Lemma 2 and Eq. 106)} \\
&\leq 4C^2\left[1 + \frac{2}{\alpha^2} + C^2(j+t_0)^{1-2\theta} B_{t,\delta}^2\right] \quad \text{(Eq. 107)} \\
&:= B'_{j,t,\delta}.
\end{aligned} \tag{110}
$$

If we use the isometry of the operator $\mathcal{L}_K^{\frac{1}{2}} : L^2_{p^\alpha} \to \mathbb{H}$, and the fact that it is a compact operator, then there exists an orthonormal eigensystem $(\mu_k, \phi_k)_{k\in\mathbb{N}}$ of $\mathcal{L}_K$, where $\{\mu_k\}_{k\in\mathbb{N}}$ are strictly positive and arranged in decreasing order (see Proposition 2.2 in Dieuleveut (2017)).

Let us define $a_i = \eta_i\lambda_i + \eta_i\mu_j$. First notice that for $j \leq t$, given Eq 42:

$$
\left\|\bar{\Pi}_{j+1}^t\right\| \leq \left\|\prod_{i=j+1}^t (I_{\mathbb{H}} - \eta_i \mathbf{A}_i)\right\| \leq \prod_{i=j+1}^t (1 - \eta_i(\lambda_i + \mu_j)) = \prod_{i=j+1}^t (1 - a_i). \tag{111}
$$

Then, we can verify the following inequality:

$$
\begin{aligned}
\sum_{j=1}^t \mathbb{E}\left[\|\Upsilon_j\|_{L^2_{p^\alpha}}^2 \mid \Xi_{j-1}\right] &= \sum_{j=1}^t \mathbb{E}\left[\left\|\mathcal{L}_K^{\frac{1}{2}}\Upsilon_j\right\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] = \sum_{j=1}^t \eta_j^2 \mathbb{E}\left[\left\|\mathcal{L}_K^{\frac{1}{2}}\bar{\Pi}_{j+1}^t \epsilon_j\right\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \\
&= \sum_{j=1}^t \left(\eta_j^2 \left\|\bar{\Pi}_{j+1}^t \mathcal{L}_K \bar{\Pi}_{j+1}^t\right\|\right) \mathbb{E}\left[\|\epsilon_j\|_{\mathbb{H}}^2 \mid \Xi_{j-1}\right] \leq \sum_{j=1}^t \eta_j^2 B'_{j,t,\delta} \left\|\bar{\Pi}_{j+1}^t \mathcal{L}_K \bar{\Pi}_{j+1}^t\right\| \quad \text{(Eq. 110)} \\
&\leq \sup_{\{\mu_k : k\in\mathbb{N}\}} \sum_{j=1}^t \eta_j^2 B'_{j,t,\delta}\mu_k \prod_{i=j+1}^t (1-a_i)^2 \quad \text{(Eq. 111)} \\
&= \sup_{\{\mu_k : k\in\mathbb{N}\}} \left[\sup_j \eta_j B'_{j,t,\delta} \prod_{i=j+1}^t (1-a_i)\right]\left[\sum_{j=1}^t \eta_j\mu_k \prod_{i=j+1}^t (1-a_i)\right].
\end{aligned} \tag{112}
$$

For a large value of t_0 we can verify for the first element of the product:

$$
\begin{aligned}
\sup_j \eta_j B'_{j,t,\delta} \prod_{i=j+1}^t (1-a_i) &\leq \sup_j \eta_j B'_{j,t,\delta} \prod_{i=j+1}^t (1-\eta_i\lambda_i) \\
&\leq \sup_j \eta_j B'_{j,t,\delta} \prod_{i=j+1}^t (1-\eta_i\lambda_i) \\
&\leq 4aC^2 \sup_j \frac{j+t_0}{\bar{t}} \left(\frac{1+\frac{2}{\alpha^2}}{(j+t_0)^\theta} + \frac{C^2 B_{t,\delta}^2}{(j+t_0)^{3\theta-1}}\right) \\
&\leq \frac{4aC^2}{\bar{t}^\theta}\left(1 + \frac{2}{\alpha^2} + \frac{C^2 B_{t,\delta}^2}{\bar{t}^{(2\theta-1)}}\right).
\end{aligned} \tag{113}
$$

For the second element of the product, we can verify:

$$
\sum_{j=1}^t \eta_j\mu_k \prod_{i=j+1}^t (1-a_i) \leq \sum_{j=1}^t (1-(1-\eta_j\mu_k)) \prod_{i=j+1}^t (1-\eta_i\mu_k) = 1 - \prod_{i=1}^t (1-\eta_i\mu_k) \leq 1. \tag{114}
$$

By combining both bounds 113 and 114 we obtain:

$$\sum_{j=1}^{t} \mathbb{E}\left[\|\Upsilon_j\|^2_{L^2_{p^\alpha}} \mid \Xi_{j-1}\right] \leq \frac{4aC^2}{\bar{t}^\theta}\left(1+\frac{2}{\alpha^2}+\frac{C^2B^2_{t,\delta}}{\bar{t}^{2\theta-1}}\right) \tag{115}$$

Now we will identify M. Let us start by upperbounding the following term via Lemma 2:

$$\begin{aligned}
\left\|K(x'_j,\cdot)-LR_j(f_{j-1})\right\|_{\mathbb{H}} &= \left\|K(x'_j,\cdot)-LR_j(r^\alpha+g_{j-1}+h_{j-1})\right\|_{\mathbb{H}} \\
&\leq \left\|K(x'_j,\cdot)\right\|_{\mathbb{H}}+\|LR_j(r^\alpha+g_{j-1})\|_{\mathbb{H}}+\|LR_j(h_{j-1})\|_{\mathbb{H}} \\
&\leq C+\|LR_j\|\,\|(r^\alpha+g_{j-1})\|_{\mathbb{H}}+\|LR_j\|\,\|h_{j-1}\|_{\mathbb{H}} \\
&\leq C+\frac{3C^2}{\alpha\sqrt{\lambda_{j-1}}}+C^2B_{t,\delta}(j+t_0)^{\frac{1}{2}-\theta} \qquad \left(\text{Lemma 2, Assumption 2 and Eq. 107}\right) \\
&= C_{j,t,\delta}.
\end{aligned} \tag{116}$$

By using the fact that f_{j-1} is Ξ_{j-1} measurable we have:

$$\begin{aligned}
\epsilon_j &= (\mathcal{L}_K-LR_j)f_{j-1}+K(x'_j,\cdot)-\mathcal{L}_K r^\alpha \\
&= (\mathcal{L}_K-LR_j)f_{j-1}+K(x'_j,\cdot)-\mathbb{E}_{p^\alpha(y)}[K(y,\cdot)r^\alpha(y)] \qquad \left(\text{Expr. 19}\right) \\
&= K(x'_j,\cdot)-LR_j(f_{j-1})-\mathbb{E}\left[K(x'_j,\cdot)-LR_j(f_{j-1})\,|\,\Xi_{j-1}\right].
\end{aligned} \tag{117}$$

where the last inequality is due to the definition of the likelihood-ratio and the fact that the observations are independent in time.

We upperbound the norm $\|\epsilon_j\|_{\mathbb{H}}$ by:

$$\begin{aligned}
\|\epsilon_j\|_{\mathbb{H}} &= \left\|K(x'_j,\cdot)-LR_j(f_{j-1})-\mathbb{E}\left[K(x'_j,\cdot)-LR_j(f_{j-1})\,|\,\Xi_{j-1}\right]\right\|_{\mathbb{H}} \qquad \left(\text{Eq. 117}\right) \\
&\leq \left\|K(x'_j,\cdot)-LR_j(f_{j-1})\right\|_{\mathbb{H}}+\left\|\mathbb{E}\left[K(x'_j,\cdot)-LR_j(f_{j-1})\,|\,\Xi_{j-1}\right]\right\|_{\mathbb{H}} \\
&\leq \left\|K(x'_j,\cdot)-LR_j(f_{j-1})\right\|_{\mathbb{H}}+\mathbb{E}\left[\left\|K(x'_j,\cdot)-LR_j(f_{j-1})\right\|_{\mathbb{H}}\,|\,\Xi_{j-1}\right] \qquad \left(\text{Jensen's inequality}\right) \\
&\leq 2\left(C+\frac{3C^2}{\sqrt{\lambda_{j-1}}}+C^2B_{t,\delta}(j+t_0)^{\frac{1}{2}-\theta}\right) \qquad \left(\text{Eq. 116}\right)=2C_{j,t,\delta}.
\end{aligned} \tag{118}$$

Therefore by using the hypothesis $t_0^\theta \geq 2+4C^2a$, we can deduce $\frac{C\sqrt{a}}{\bar{t}^{\frac{\theta}{2}}} \leq 1$ and:

$$\begin{aligned}
\|\Upsilon_j\|_{L^2_{p^\alpha}} &= \left\|\mathcal{L}_K^{\frac{1}{2}}\Upsilon_j\right\|_{\mathbb{H}} \leq \eta_j\left\|\mathcal{L}_K^{\frac{1}{2}}\Pi^t_{j+1}\epsilon_j\right\|_{\mathbb{H}} \leq 2\eta_j C_{j,t,\delta}\left\|\Pi^t_{j+1}\mathcal{L}_K\Pi^t_{j+1}\right\|^{\frac{1}{2}} \qquad \left(\text{Eq. 118}\right) \\
&\leq 2C\sup_j \eta_j C_{j,t,\delta}\prod_{i=j+1}^{t}(1-\eta_j\lambda_j) \qquad \left(\text{As } \|\mathcal{L}_K\|\leq C^2\right) \\
&= 2aC^2\sup_j \frac{j+t_0}{\bar{t}}\left(\frac{1}{(j+t_0)^\theta}+\frac{3C\sqrt{a}}{\alpha(j+t_0)^{\left(\frac{3\theta-1}{2}\right)}}+\frac{CB_{t,\delta}}{(j+t_0)^{2\theta-\frac{1}{2}}}\right) \qquad \left(\text{Eq. 116}\right) \\
&\leq 2aC^2\left(\frac{1}{\bar{t}^\theta}+\frac{3C\sqrt{a}}{\alpha\bar{t}^{\left(\frac{3\theta-1}{2}\right)}}+\frac{CB_{t,\delta}}{\bar{t}^{2\theta-\frac{1}{2}}}\right) \\
&\leq \frac{2aC^2}{\bar{t}^\theta}\left(1+\frac{3C\sqrt{a}}{\alpha\bar{t}^{\left(\frac{\theta-1}{2}\right)}}+\frac{CB_{t,\delta}}{\bar{t}^{\theta-\frac{1}{2}}}\right) \\
&\leq \frac{2\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(1+\frac{3C^2a}{\alpha\bar{t}^{\left(\theta-\frac{1}{2}\right)}}+\frac{CB_{t,\delta}}{\bar{t}^{\theta-\frac{1}{2}}}\right) \qquad \left(\text{The hypothesis } t_0^\theta\geq 2+4C^2a \text{ implies } \frac{t_0^{\frac{\theta}{2}}}{C\sqrt{a}}\geq 1\right)
\end{aligned} \tag{119}$$

Both Inequalities 115 and 119 imply that the hypothesis of Proposition A.3 is satisfied. Then the inequality holds with a probability at least $1-\delta$ for $\delta \in (0,1)$ we have:

$$\begin{aligned}
\sup_{1\leq k\leq t}\left\|\sum_{j=1}^{k}\Upsilon_j\right\|_{L^2_{p^\alpha}} &\leq 4\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1}{3}+\frac{C^2a}{\alpha\bar{t}^{(\theta-\frac{1}{2})}}+\frac{CB_{t,\delta}}{3\bar{t}^{(\theta-\frac{1}{2})}}+1+\frac{\sqrt{2}}{\alpha}+\frac{CB_{t,\delta}}{\bar{t}^{(\theta-\frac{1}{2})}}\right)\log\left(\frac{2}{\delta}\right)\\
&= 4\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{4}{3}+\frac{\sqrt{2}}{\alpha}+\left(\frac{4C}{3}\right)\frac{B_{t,\delta}}{\bar{t}^{(\theta-\frac{1}{2})}}+\frac{C^2a}{\alpha\bar{t}^{(\theta-\frac{1}{2})}}\right)\log\left(\frac{2}{\delta}\right)\\
&\leq 8\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1+\alpha}{\alpha}\right)\log\left(\frac{2}{\delta}\right)\\
&+\left(\frac{16C^3}{3\alpha}\right)\left[5a^{\frac{5}{2}}+a^{\frac{3}{2}}(14Ca^2+18)\sqrt{\log(\bar{t})}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}+\left(\frac{4a^{\frac{3}{2}}C^3}{\alpha}\right)\frac{\log\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&\leq 8\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1+\alpha}{\alpha}\right)\log\left(\frac{2}{\delta}\right)+\left(\frac{16C^3}{3\alpha}\right)\left[5a^{\frac{5}{2}}+a^{\frac{7}{2}}(14C+2)\sqrt{\log(\bar{t})}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&+\left(\frac{4a^{\frac{7}{2}}C^3}{\alpha}\right)\frac{\log\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\qquad \left(\text{The hypothesis } a\geq 4\right)\\
&\leq 8\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1+\alpha}{\alpha}\right)\log\left(\frac{2}{\delta}\right)+\left[\frac{32a^{\frac{5}{2}}C^3}{\alpha}+\frac{4a^{\frac{7}{2}}C^3}{\alpha}(20C+4)\sqrt{\log(\bar{t})}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&+\left(\frac{4a^{\frac{7}{2}}C^3}{\alpha}\right)\frac{\log\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&\leq 8\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1+\alpha}{\alpha}\right)\log\left(\frac{2}{\delta}\right)+\left[\frac{32a^{\frac{5}{2}}C^3}{\alpha}+\frac{4a^{\frac{7}{2}}C^3}{\alpha}(20C+4+\frac{1}{\log(2)})\sqrt{\log(\bar{t})}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&\leq 8\frac{\sqrt{a}C}{\bar{t}^{\frac{\theta}{2}}}\left(\frac{1+\alpha}{\alpha}\right)\log\left(\frac{2}{\delta}\right)+\left[\frac{32a^{\frac{5}{2}}C^3}{\alpha}+\frac{8a^{\frac{7}{2}}C^3}{\alpha}(10C+3)\sqrt{\log(\bar{t})}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}\\
&\leq \frac{\sqrt{a}B_4}{\bar{t}^{\frac{\theta}{2}}}\log\left(\frac{2}{\delta}\right)+\left[B_5a^{\frac{5}{2}}+B_6a^{\frac{7}{2}}\sqrt{\log\bar{t}}\right]\frac{\log^2\left(\frac{2}{\delta}\right)}{\bar{t}^{\frac{3\theta-1}{2}}}.
\end{aligned}$$

Where we have used the assumption $t_0\geq 3$ and the constants B_4, B_5, B_6.

$$8C\left(\frac{1+\alpha}{\alpha}\right)\leq\frac{16C}{\alpha}=B_4\quad B_5=\frac{32C^3}{\alpha}\quad B_6=\frac{8C^3(10C+3)}{\alpha}.$$

□

Upperbound for $\mathcal{E}'_{sample}(t)=\left\|\sum_{j=1}^{t}\eta_j\Pi_{j+1}^{t}(A_jf_{\lambda_j}-b_j)\right\|_{\mathbb{H}}$**.**

In this section, we focus on developing the required components for proving Theorem 10.

Lemma 8. *We have:*

1. $\|A_tf_{\lambda_t}-b_t\|_{\mathbb{H}}\leq\frac{1}{\sqrt{\lambda_t}}\left(\frac{C+1}{\alpha}+C\right)$, *if* $t_0^{1-\theta}\geq b$;
2. $\mathbb{E}\left[\|A_tf_{\lambda_t}-b_t\|^2_{\mathbb{H}}\right]\leq 2C^2\left(\frac{1+\alpha^2}{\alpha^2}\right)$.

Proof. By the definition given of f_{λ_t} in Eq. 37, we have that for any $\lambda>0$:

$$\mathbb{E}_{p^\alpha(y)}[(f_\lambda-r^\alpha)^2(y)]+\lambda\|f_\lambda\|^2_{\mathbb{H}}\leq\mathbb{E}_{p^\alpha(y)}[(r^\alpha)^2(y)]\leq\frac{1}{\alpha^2},\tag{120}$$

which implies:

$$\|f_\lambda\|_{\mathbb{H}}\leq\frac{1}{\alpha\sqrt{\lambda}}.\tag{121}$$

On the other hand, using the close-form solution of f_λ we get:

$$\|f_\lambda\|_{\mathcal{L}^2_{p^\alpha}} = \left\|(\mathcal{L}_K + \lambda I)^{(-1)}\mathcal{L}_K r^\alpha\right\|_{\mathcal{L}^2_{p^\alpha}} \leq \left\|(\mathcal{L}_K + \lambda I)^{(-1)}\mathcal{L}_K\right\| \|r^\alpha\|_{\mathcal{L}^2_{p^\alpha}} \leq \frac{1}{\alpha}. \tag{122}$$

Moreover, we know:

$$\begin{aligned}
\mathbb{E}\left[\|LR_t f_{\lambda_t}\|^2_{\mathbb{H}}\right] &= \mathbb{E}\left[\langle(1-\alpha)f_{\lambda_t}(x_t)K(x_t,\cdot) + \alpha f_{\lambda_t}(x'_t)K(x'_t,\cdot), (1-\alpha)f_{\lambda_t}(x_t)K(x_t,\cdot) + \alpha f_{\lambda_t}(x'_t)K(x'_t,\cdot)\rangle_{\mathbb{H}}\right] \\
&= \mathbb{E}\left[(1-\alpha)^2 K(x_t,x_t)f^2_{\lambda_t}(x_t) + 2(1-\alpha)\alpha K(x_t,x'_t)f_{\lambda_t}(x_t)f_{\lambda_t}(x'_t) + \alpha^2 K(x'_t,x'_t)f^2_{\lambda_t}(x'_t)\right] \\
&\leq C^2\mathbb{E}\left[((1-\alpha)f_{\lambda_t}(x_t) + \alpha f_{\lambda_t}(x'_t))^2\right] \quad (\text{Assumption } 2) \\
&\leq C^2\mathbb{E}\left[(1-\alpha)f^2_{\lambda_t}(x_t) + \alpha f^2_{\lambda_t}(x'_t)\right] \quad (\text{Jensen's inequality}) \\
&= C^2 \|f_{\lambda_t}\|^2_{L^2_{p^\alpha}} \leq \frac{C^2}{\alpha^2}.
\end{aligned} \tag{123}$$

By putting these elements together, we can proove the first point of Lemma 8:

$$\begin{aligned}
\|A_t f_{\lambda_t} - b_t\|_{\mathbb{H}} &\leq (1-\alpha)\|K(x_t,\cdot)f_{\lambda_t}(x_t)\|_{\mathbb{H}} + \alpha\|K(x'_t,\cdot)f_{\lambda_t}(x'_t)\|_{\mathbb{H}} + \lambda_t\|f_{\lambda_t}\|_{\mathbb{H}} + \|K(x'_t,\cdot)\|_{\mathbb{H}} \\
&\leq C\left(\frac{1}{\alpha\sqrt{\lambda_t}} + 1\right) + \frac{\sqrt{\lambda_t}}{\alpha} \quad (\text{Eq. } 121) \\
&\leq \frac{1}{\sqrt{\lambda_t}}\left(\frac{C}{\alpha} + C\sqrt{\lambda_t} + \frac{\lambda_t}{\alpha}\right) \\
&\leq \frac{1}{\sqrt{\lambda_t}}\left(\frac{C+1}{\alpha} + C\right),
\end{aligned}$$

where in the last equality we have used the hypothesis $t_0^{1-\theta} \geq b$, which implies $\lambda_t \leq 1$.

Given the definition of f_λ, we have $\mathcal{L}_K f_\lambda + \lambda f_\lambda = \mathcal{L}_K r^\alpha$, which leads to:

$$A_t f_{\lambda_t} - b_t = LR_t f_{\lambda_t} - K(x'_t,\cdot) + \lambda_t f_{\lambda_t} = (LR_t - \mathcal{L}_K)f_{\lambda_t} + \mathcal{L}_K r^\alpha - K(x'_t,\cdot), \tag{124}$$

Then, we verify the second point of Lemma 8:

$$\begin{aligned}
\mathbb{E}\left[\|A_t f_{\lambda_t} - b_t\|^2_{\mathbb{H}}\right] &= \mathbb{E}\left[\|(LR_t - \mathcal{L}_K)f_{\lambda_t} + \mathcal{L}_K r^\alpha - K(x'_t,\cdot)\|^2_{\mathbb{H}}\right] \\
&\leq 2\left[\mathbb{E}\left[\|(LR_t - \mathcal{L}_K)f_{\lambda_t}\|^2_{\mathbb{H}} + \|\mathcal{L}_K r^\alpha - K(x'_t,\cdot)\|^2_{\mathbb{H}}\right]\right] \quad \left(2\langle a,b\rangle_{\mathbb{H}} \leq \|a\|^2_{\mathbb{H}} + \|b\|^2_{\mathbb{H}}\right) \\
&\leq 2\left[\mathbb{E}\left[\|LR_t f_{\lambda_t}\|^2_{\mathbb{H}} + \|K(x'_t,\cdot)\|^2_{\mathbb{H}}\right]\right] \quad (\text{After developing the norm and taking expectations}) \\
&\leq 2C^2\left(\frac{1+\alpha^2}{\alpha^2}\right) \quad (\text{Eq. } 123 \text{ and Assumption } 2).
\end{aligned}$$

□

Proof Theorem 10.

Proof. The idea of the proof is to use Lemma 9 to generate a probabilistic bound for the quantity:

$$\mathcal{E}'_{\text{sample}}(t) = \left\|\sum_{j=1}^{t}\eta_j\Pi^t_{j+1}(A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}}.$$

We have shown in Sec. A.4 that the process $\{\eta_j\Pi^t_{j+1}(A_j f_{\lambda_j} - b_j)\}^t_{j=1}$ is a reversed martingale difference with respect to the sequence of sigma algebras $\mathcal{B}_j = \sigma((x_j,x'_j),...,(x_t,x'_t),...)$. The only element to finish the proof is to identify M and σ^2.

Given the definition of the random variables $\{A_t\}_{t\in\mathbb{N}}$, then for $t > 1$ we can verify A_t is a positive linear operator:

$$\langle A_t f, f\rangle_{\mathbb{H}} = (1-\alpha)f^2(x_t) + \alpha f^2(x'_t) + \lambda_t\|f\|^2_{\mathbb{H}} \geq 0 \quad \text{for } f \in \mathbb{H}. \tag{125}$$

Moreover as $\|A_t\| \geq \lambda_t$ we have:

$$\|I_{\mathbb{H}} - \eta_t A_t\| \leq (1 - \eta_t \lambda_t). \tag{126}$$

Let us consider the following group of expressions:

$$\begin{aligned}
\eta_j \left\|\Pi_{j+1}^t\right\| &= \eta_j \prod_{i=j+1}^{t} (1 - \eta_i \lambda_i) \leq \eta_j \exp\left(-\sum_{i=j+1}^{t} \eta_i \lambda_i\right) \\
&= \eta_j \exp\left(-\sum_{i=j+1}^{t} \frac{ab}{t_0 + i}\right) \\
&\leq \eta_j \exp\left(-ab \log\left(\frac{\bar{t}}{t_0 + j + 1}\right)\right) \\
&= \frac{a}{(j + t_0)^\theta} \left(\frac{t_0 + j + 1}{\bar{t}}\right)^{ab} \\
&= a \frac{(t_0 + j)^{ab - \theta}}{\bar{t}^{ab}} (1 + \frac{1}{t_0 + j})^{ab} \\
&\leq a \frac{(t_0 + j)^{ab - \theta}}{\bar{t}^{ab}} (1 + \frac{1}{t_0})^{ab} \\
&\leq \frac{ea(t_0 + j)^{ab - \theta}}{\bar{t}^{ab}}.
\end{aligned} \tag{127}$$

This implies:

$$\begin{aligned}
\mathbb{E}\left[\|\zeta_j\|_{\mathbb{H}}^2 \mid \mathcal{B}_{j+1}\right] &= \mathbb{E}\left[\left\|\eta_j \Pi_{j+1}^t (A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}}^2 \mid \mathcal{B}_{j+1}\right] \\
&\leq \eta_j^2 \left\|\Pi_{j+1}^t\right\|^2 \mathbb{E}\left[\left\|A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}}^2 \mid \mathcal{B}_{j+1}\right] \\
&\leq \frac{2(eaC)^2 (t_0 + j)^{2ab - 2\theta}}{\bar{t}^{2ab}} \left(\frac{1 + \alpha^2}{\alpha^2}\right) \qquad \big(\text{Lemma 8, Eq. 127 and Assumptions 1, 2}\big).
\end{aligned} \tag{128}$$

If $t_0 \geq 2$, we will find:

$$\begin{aligned}
\sum_{j=1}^{t} \mathbb{E}\left[\|\zeta_j\|_{\mathbb{H}}^2 \mid \mathcal{B}_{j+1}\right] &\leq \sum_{j=1}^{t} \frac{2(eaC)^2 (t_0 + j)^{2ab - 2\theta}}{\bar{t}^{2ab}} \left(\frac{1 + \alpha^2}{\alpha^2}\right) \\
&\leq \frac{2(eaC)^2}{\bar{t}^{2ab}} \left(\frac{1 + \alpha^2}{\alpha^2}\right) \int_1^t (t_0 + s)^{2ab - 2\theta} ds \\
&\leq \frac{2(eaC)^2}{\bar{t}^{2ab}} \left(\frac{1 + \alpha^2}{\alpha^2}\right) \left(\frac{1}{2ab - 2\theta + 1}\right) \bar{t}^{2ab - 2\theta + 1} \\
&= \begin{cases} \frac{(eaC)^2}{ab - \theta + \frac{1}{2}} \bar{t}^{-2\theta + 1} & , \text{ if } \quad ab > \theta - \frac{1}{2} \\ \frac{(eaC)^2}{\theta - \frac{1}{2} - ab} \bar{t}^{-2ab} & , \text{ if } \quad ab < \theta - \frac{1}{2} \end{cases} \\
&\leq \frac{(eaC)^2}{\left|ab - \theta + \frac{1}{2}\right|} \bar{t}^{-2\left(ab \wedge (\theta - \frac{1}{2})\right)}
\end{aligned} \tag{129}$$

On the other hand, if $t_0^{1-\theta} \geq b$ we have:

$$\begin{aligned}
\|\zeta_j\|_{\mathbb{H}} &= \left\|\eta_j \Pi_{j+1}^t (A_j f_{\lambda_j} - b_j)\right\|_{\mathbb{H}} \\
&\leq \frac{1}{\sqrt{\lambda_t}} \frac{ea(t_0+j)^{ab-\theta}}{\bar{t}^{ab}} \left(\frac{C+1}{\alpha} + C\right) \qquad \left(\text{Lemma 8 and Eq. 127}\right) \\
&= \frac{ea(t_0+j)^{ab-\frac{(3\theta+1)}{2}}}{\sqrt{b}\bar{t}^{ab}} \left(\frac{C+1}{\alpha} + C\right) \\
&= \begin{cases} \frac{ea}{\sqrt{b}} \left(\frac{C+1}{\alpha} + C\right) \bar{t}^{-\frac{(3\theta+1)}{2}} & , \text{ if } \quad ab > \frac{3\theta-1}{2} \\ \frac{ea}{\sqrt{b}} \left(\frac{C+1}{\alpha} + C\right) \bar{t}^{-2ab} & , \text{ if } \quad ab < \frac{3\theta-1}{2} \end{cases} \\
&= \frac{ea}{\sqrt{b}} \left(\frac{C+1}{\alpha} + C\right) \bar{t}^{-\left(ab \wedge \frac{3\theta-1}{2}\right)}
\end{aligned} \tag{130}$$

Then by Lemma 9 we get with probability $1-\delta$:

$$\begin{aligned}
\mathcal{E}'_{\text{sample}}(t) &\leq 2\left(\frac{ea}{3\sqrt{b}} \left(\frac{C+1}{\alpha} + C\right) \bar{t}^{-\left(ab \wedge \frac{3\theta-1}{2}\right)} + \sqrt{\frac{1}{|ab-\theta+\frac{1}{2}|}} eaC\bar{t}^{-\left(ab\wedge(\theta-\frac{1}{2})\right)}\right) \log\left(\frac{2}{\delta}\right) \\
&= ab^{-\frac{1}{2}} B'_4 \bar{t}^{-\left(ab \wedge \frac{3\theta-1}{2}\right)} + B'_5 a\bar{t}^{-\left(ab\wedge(\theta-\frac{1}{2})\right)},
\end{aligned} \tag{131}$$

where

$$B'_4 = \frac{2e}{3}\left(\frac{C+1}{\alpha} + C\right) \log\left(\frac{2}{\delta}\right) \qquad B'_5 = 2\sqrt{\frac{1}{|ab-\theta+\frac{1}{2}|}} eC \log\left(\frac{2}{\delta}\right). \tag{132}$$

□

B AUXILIARY RESULTS

The following result is frequently used in Appendix A. It was first proved by the Proposition A.3 in Tarrès and Yao (2014). We include the result for completeness.

Lemma 9. ***(Proposition A.3 (Pinelis-Bernstein)*** *Tarrès and Yao (2014)) Let ζ_i be a martingale difference sequence in a Hilbert space. Suppose that almost surely $\|\zeta_i\| \leq M$ and $\sum_{i=1}^t \mathbb{E}\left[\|\zeta_i\|_{\mathbb{H}}^2 \,|\, \Xi_{i-1}\right] \leq \sigma_t^2$. Then the following holds with probability at least $1-\delta$ (with $\delta \in (0,1)$),*

$$\sup_{1\leq k\leq t} \left\|\sum_{i=1}^k \zeta_i\right\| \leq 2\left(\frac{M}{3} + \sigma_t\right) \log(\frac{2}{\delta}).$$

The last inequality can be as well be applied for ζ_i being a reversed martingales difference sequence in a Hilbert space. With a small change where $\sum_{i=1}^t \mathbb{E}\left[\|\zeta_i\|_{\mathbb{H}}^2 \,|\, \Xi_{i-1}\right] \leq \sigma_t^2$ is replaced by $\sum_{i=1}^t \mathbb{E}\left[\|\zeta_i\|_{\mathbb{H}}^2 \,|\, \mathcal{B}_{i+1}\right] \leq \sigma_t^2$, where $\mathcal{B}_{i+1}$ is the sigma-algebra generated by observations after index i.