1

Projection of fBm on the Space of Martingales

Consider the fractional Brownian motion (fBm) with Hurst index $H \in (0,1)$. Its definition and properties will be considered in more detail in section 1.1; however, let us mention immediately that fBm is a Gaussian process and anyhow not a martingale or even a semimartingale for $H \neq \frac{1}{2}$. Hence, a natural question arises: what is the distance between fBm and the space of Gaussian martingales in an appropriate metric and how do we determine the projection of fBm on the space of Gaussian martingales? Why is it not reasonable to consider non-Gaussian martingales? In this chapter, we will answer this and other related questions. The chapter is organized as follows. In section 1.1, we give the main properties of fBm, including its integral representations. In section 1.2, we formulate the minimizing problem simplifying it at the same time. In section 1.3, we strictly propose a positive lower bound for the distance between fBm and the space of Gaussian martingales. Sections 1.4 and 1.5 are devoted to the general problem of minimization of the functional f on $L_2([0,1])$ that has the following form:

$$f(x) = \sup_{t\in[0,1]} \left(\int_0^t \big(z(t,s) - x(s)\big)^2 \, ds \right)^{1/2} \qquad [1.1]$$

with arbitrary kernel $z(t,s)$ satisfying condition

(A) for any $t \in [0,1]$ the kernel $z(t,\cdot) \in L_2([0,t])$ and

$$\sup_{t\in[0,1]} \int_0^t z(t,s)^2 \, ds < \infty. \qquad [1.2]$$

We shall call the functional f the *principal functional*. It is proved in section 1.4 that the principal functional f is convex, continuous and unbounded on infinity, consequently the minimum is reached. Section 1.5 gives an example of the kernel $z(t,s)$ where a minimizing function for the principal functional is not unique (moreover, being convex, the set of minimizing functions is infinite). Sections 1.6–1.8 are devoted to the problem of minimization of principal functional f with the kernel z corresponding to fBm, i.e. with the kernel z from [1.7]. It is proved in section 1.6 that in this case, the minimizing function for the principal functional is unique. In section 1.7 it is proved that the minimizing function has a special form, namely a probabilistic representation, and many properties of the minimizing function have been established. Since we have no explicit analytical representation of the minimizing function, in section 1.8 we provide the discrete-time counterpart of the minimization problem and give the results explaining how to calculate the minimizing function numerically via evaluation of the Chebyshev center, illustrating the numerics with a couple of plots.

1.1. fBm and its integral representations

In this section, we define fBm and collect some of its main properties. We refer to the books [BIA 08, MIS 08, MIS 18, NOU 12] for the detailed presentation of this topic.

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the standard assumptions.

DEFINITION 1.1.– *An fBm with associated Hurst index* $H \in (0,1)$ *is a Gaussian process* $B^H = \{B_t^H, \mathcal{F}_t, t \geq 0\}$, *such that*

1) $\mathrm{E}B_t^H = 0$, $t \geq 0$,

2) $\mathrm{E}B_t^H B_s^H = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right)$, $s,t \geq 0$.

The following statements can be derived directly from the above definition.

1) If $H = \frac{1}{2}$, then an fBm is a standard Wiener process.

2) An fBm is self-similar with the self-similarity parameter H, i.e. $\left\{B_{ct}^H\right\} \stackrel{d}{=} \left\{c^H B_t^H\right\}$ for any $c > 0$. Here, $\stackrel{d}{=}$ means that all finite-dimensional distributions of both processes coincide.

3) An fBm has stationary increments that is implied by the form of its incremental covariance:

$$\mathrm{E}\left(B_t^H - B_s^H\right)^2 = (t-s)^{2H}. \qquad [1.3]$$

4) The increments of an fBm are independent only in the case $H = 1/2$. They are negatively correlated for $H \in (0, 1/2)$ and positively correlated for $H \in (1/2, 1)$.

Due to the Kolmogorov continuity theorem, property [1.3] implies that an fBm has a continuous modification. Moreover, this modification is γ-Hölder continuous on each finite interval for any $\gamma \in (0, H)$.

It is also well-known that an fBm is not a process of bounded variation. If $H \neq \frac{1}{2}$, then it is neither a semimartingale nor a Markov process.

An fBm can be represented as an integral of a deterministic kernel with respect to the standard Wiener process in several ways.

We start with the *Molchan representation* (or Volterra-type representation) of fBm $B^H = \{B_t^H, \mathcal{F}_t, t \geq 0\}$ via the Wiener process on a finite interval (see, for example, [NOR 99b, NUA 03]). It states that a Wiener process $W = \{W_t, \mathcal{F}_t, t \geq 0\}$ exists, such that for any $t \geq 0$

$$B_t^H = \int_0^t z(t,s)\, dW_s, \qquad [1.4]$$

where the Molchan kernel is defined by

$$z(t,s) = c_H \biggl(t^{H-1/2} s^{1/2-H} (t-s)^{H-1/2} - \left(H - \tfrac{1}{2}\right) s^{1/2-H} \int_s^t u^{H-3/2} (u-s)^{H-1/2}\, du \biggr) \mathbf{1}_{0<s<t}, \qquad [1.5]$$

with

$$c_H = \left(\frac{2H\Gamma(\frac{3}{2} - H)}{\Gamma(H + \frac{1}{2})\Gamma(2 - 2H)} \right)^{1/2}. \qquad [1.6]$$

In the case $H \in \left(\frac{1}{2}, 1\right)$, the kernel $z(t,s)$ can be simplified to

$$z(t,s) = c_H \left(H - \tfrac{1}{2}\right) s^{1/2-H} \int_s^t u^{H-1/2} (u-s)^{H-3/2}\, du\, \mathbf{1}_{0<s<t}. \qquad [1.7]$$

The *Mandelbrot–Van Ness representation*, or moving-average representation, was obtained in [MAN 68]. It states that an fBm B^H can be represented as

$$B_t^H = \int_{-\infty}^t z_1(t,s)\, d\widetilde{W}_s, \qquad [1.8]$$

where $\widetilde{W}$ is a two-sided Wiener process, and the Volterra kernel z_1 is defined by the formula

$$z_1(t,s) = c_H \left((t-s)_+^{H-1/2} - (-s)_+^{H-1/2} \right), \qquad [1.9]$$

and $x_+ = \max\{x, 0\}$. This representation defines fBm on the whole axis, but in what follows we shall consider only the processes that are defined on $\mathbb{R}^+$. The next result demonstrates that the finite-dimensional distributions of fBm are of non-degenerate form.

THEOREM 1.1.– *Let* $0 < H < 1$, $B^H = \{B_t^H, t \geq 0\}$ *be an fBm with associated Hurst index* H. *Then, the finite-dimensional distributions of* B^H *have a non-singular covariance matrix, i.e. for any set of points* $t_1, \ldots, t_n$, $0 < t_1 < \ldots < t_n$, *the covariance matrix* $\mathrm{E}vv^\top$ *of the slice-vector* $v = (B_{t_1}^H, \ldots, B_{t_n}^H)$ *is non-singular.*

PROOF.– Recall that fBm B^H admits the Mandelbrot–Van Ness representation [1.8]. Now, let $0 < t_1 < \ldots < t_n$. We carry out the proof by contradiction. Thus, assume that the vector $v = (B_{t_1}^H, \ldots, B_{t_n}^H)$, which has multivariate Gaussian distribution with zero mean, has a singular covariance matrix. Then,

$$\mathrm{E}((vv^\top) \cdot (\alpha_1, \ldots, \alpha_n)^\top) = 0$$

for some non-zero vector $(\alpha_1, \ldots, \alpha_n)^\top$. In turn, it follows that

$$\mathrm{E}((v^\top) \cdot (\alpha_1, \ldots, \alpha_n)^\top)^2 = 0,$$

or, that is the same:

$$\mathrm{E}\left(\sum_{k=1}^{n} \alpha_k B_{t_k}^H \right)^2 = 0. \qquad [1.10]$$

Without loss of generality, we can assume that $\alpha_k \neq 0$ for all $k = 1, \ldots, n$. Denote $t_0 = 0$. With Mandelbrot–Van Ness representation [1.8], we obtain

$$\sum_{k=1}^{n} \alpha_k B_{t_k}^H = c_H \int_{-\infty}^{t_n} a(s)\, dW_s,$$

where:

$$a(s) = \textstyle\sum_{k=1}^{n} \alpha_k (t_k - s)^{H-1/2} - \sum_{k=1}^{n} \alpha_k (-s)^{H-1/2}, \quad \text{for} \quad s < 0;$$

$$a(s) = \sum_{k=m}^{n} \alpha_k (t_k - s)^{H-1/2}, \quad \text{for} \quad t_{m-1} < s < t_m, \quad m = 1, \ldots, n;$$
$$a(s) = 0, \quad \text{for} \quad s > t_n.$$

Observe that $a = a(s)$ is not an almost-everywhere (a.e.) zero function, particularly $a(s) \neq 0$ for all $s \in (t_{n-1}, t_n)$. Hence, $\int_{-\infty}^{t_n} a(s)^2 \, ds > 0$, and

$$\mathrm{E}\left(\sum_{k=1}^{n} \alpha_k B_{t_k}^H\right)^2 = c_H^2 \int_{-\infty}^{t_n} a(s)^2 \, ds > 0.$$

We got the contradiction with [1.10], whence the proof follows. □

1.2. Formulation of the main problem

Let $B^H = \{B_t^H, \mathcal{F}_t, t \in [0,T]\}$ be an fBm with Hurst index $H \in (0,1)$, restricted to the interval $[0,T]$. Recall that an fBm is neither a semimartingale nor a Markov process unless $H = 1/2$. Therefore, a simple and natural question is: how far is Brownian motion from being a martingale? That is, in a sense, we look for the distance between fBm and the space of martingales and for the projection of fBm on the space of (square integrable) martingales. Thus, initially, the problem is formulated in such a way: we are looking for a square integrable $\mathcal{F}$-martingale M that minimizes the value

$$\rho_H^2(M) := \sup_{t \in [0,T]} \mathrm{E}\left(B_t^H - M_t\right)^2,$$

and try to calculate or estimate the value itself. To proceed with the solution of this problem, we can use the Molchan representation [1.4] of the fBm B^H via the standard $\mathcal{F}$-Brownian motion $W = \{W_t, \mathcal{F}_t, t \in [0,T]\}$. We observe first that B^H and W generate the same filtration, and so according to the standard martingale representation theorem (see, for example, [DAV 05]), any square integrable $\mathcal{F}$-martingale M admits a representation

$$M_t = \int_0^t a(s) \, dW_s, \tag{1.11}$$

where a is an $\mathcal{F}$-adapted square integrable process.

Hence, we can write

$$\mathrm{E}\left(B_t^H - M_t\right)^2 = \mathrm{E}\left(\int_0^t \left(z(t,s) - a(s)\right) dW_s\right)^2$$

$$= \int_0^t \mathrm{E}\big(z(t,s) - a(s)\big)^2\, ds$$
$$= \int_0^t \big(z(t,s) - \mathrm{E}a(s)\big)^2\, ds + \int_0^t \mathrm{Var}\, a(s)\, ds.$$

Consequently, it is enough to minimize $\rho_H(M)$ over *Gaussian* martingales, i.e. those having representation [1.11] with a non-random a.

Hence, the main problem reduces to the following one: take the functional f from [1.1] and find

$$\begin{aligned}\rho^2\big(B^H, \mathcal{M}(L_2([0,T]))\big) &:= \inf_{x \in L_2([0,T])} \sup_{t\in[0,T]} \int_0^t \big(z(t,s) - x(s)\big)^2\, ds \\ &= \inf_{x\in L_2([0,T])} f(x),\end{aligned} \qquad [1.12]$$

and a minimizing element $a \in L_2([0,T])$ if the infimum is reached. Note that the expression being minimized involves neither an fBm nor a Wiener process; thus, the problem becomes purely analytic.

Note that it is natural to consider the integral in [1.11] with respect to the Wiener process W from the Molchan representation of B^H, which can be called the underlying Wiener process. Indeed, the distance $\mathrm{E}(B_t^H - M_t)^2$ increases if M_t is of the form $M_t = \int_0^t a(s)\, d\widetilde{W}_s$, where $\widetilde{W}$ is a Wiener process with a component independent of W. This fact is established in the following lemma.

LEMMA 1.1.– *Among all Wiener processes* $\left\{\widetilde{W}_t, \mathcal{F}_t, t \in [0,T]\right\}$*, the minimum in the expression* $\mathrm{E}\Big(B_t^H - \int_0^t a(s)\, d\widetilde{W}_s\Big)^2$ *is reached for*

$$\widetilde{W}_t = \int_0^t \mathrm{sgn}\, a(s)\, dW_s,$$

where W *is a Wiener process from the Molchan representation* [1.4] *of* B^H.

PROOF.– Let $\left\{\widetilde{W}_t, \mathcal{F}_t, t \in [0,T]\right\}$ be a Wiener process correlated with W so that

$$\mathrm{E}\widetilde{W}_t W_t = \rho(t).$$

Then,

$$\rho(t) - \rho(s) = \mathrm{E}\left(\widetilde{W}_t - \widetilde{W}_s\right)(W_t - W_s) + \mathrm{E}\widetilde{W}_s (W_t - W_s) + \mathrm{E}\left(\widetilde{W}_t - \widetilde{W}_s\right) W_s.$$

By conditioning with respect to $\mathcal{F}_s$, we can easily show that for all $t > s$, $\mathrm{E}\widetilde{W}_s (W_t - W_s) = \mathrm{E}\left(\widetilde{W}_t - \widetilde{W}_s\right) W_s = 0$. Hence, by the Cauchy–Schwarz inequality,

$$|\rho(t) - \rho(s)| \leq \sqrt{\mathrm{E}\left(\widetilde{W}_t - \widetilde{W}_s\right)^2 \mathrm{E}\left(W_t - W_s\right)^2} = |t - s|\,.$$

Therefore, $\rho(t) = \int_0^t \theta(s)\,ds$ with $|\theta(s)| \leq 1$, and

$$\widetilde{W}_t = \int_0^t \theta(s)\,dW_s + \int_0^t \sqrt{1 - \theta^2(s)}\,dZ_s,$$

where $Z = \{Z_t, \mathcal{F}_t, t \in [0,T]\}$ is a Wiener process independent of W. Then,

$$\begin{aligned}
&\mathrm{E}\left(B_t^H - \int_0^t a(s)\,d\widetilde{W}_s\right)^2 \\
&\quad = \mathrm{E}\left(\int_0^t z(t,s)\,dW_s - \int_0^t a(s)\theta(s)\,dW_s - \int_0^t a(s)\sqrt{1-\theta^2(s)}\,dZ_s\right)^2 \\
&\quad = \mathrm{E}\left(\int_0^t z(t,s)\,dW_s - \int_0^t a(s)\theta(s)\,dW_s\right)^2 + \int_0^t a^2(s)\left(1 - \theta^2(s)\right)\,ds \\
&\quad = \int_0^t \left(z(t,s) - a(s)\theta(s)\right)^2 ds + \int_0^t a^2(s)\left(1-\theta^2(s)\right)\,ds \\
&\quad = t^{2H} - 2\int_0^t a(s)\theta(s)z(t,s)\,ds + \int_0^t a^2(s)\,ds. \qquad [1.13]
\end{aligned}$$

Since $z(t,\,s) > 0$, we see that the minimum in [1.13] is reached at the function

$$a(s)\theta(s) = |a(s)|,$$

i.e. $\theta(s) = \operatorname{sgn} a(s)$. □

COROLLARY 1.1.– *Since the proof of Lemma 1.1 is valid for any non-negative kernel z, satisfying condition* [1.2], *from now on, considering the non-negative kernel z, we can restrict ourselves to Gaussian martingales of the form $M_t = \int_0^t a(s)\,dW_s$, where W is the underlying Wiener process, and function a is non-negative.*

1.3. The lower bound for the distance between fBm and Gaussian martingales

Denote $\mathcal{M}(\mathcal{K})$ the space of the Gaussian martingales of the form $M_t = \int_0^t a(s)dW_s$, where $a \in \mathcal{K} \subset L_2([0,T])$. In the following theorem, using stochastic considerations, we establish a non-zero lower bound for the distance between fBm with the Molchan kernel and the space of all Gaussian martingales, i.e. the space $\mathcal{M}(L_2([0,T]))$. All processes are considered on the fixed interval $[0,T]$.

THEOREM 1.2.– *The value*

$$\rho_T := \rho^2\big(B^H, \mathcal{M}(L_2([0,T]))\big) = \inf_{a\in L_2([0,T])} \sup_{0\le t\le T} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2$$

admits the following lower bound:

$$\rho_T \ge \max_{0\le t\le 1} \frac{\big(1-t^{2H}-(1-t)^{2H}\big)^2}{16t^{2H}} \cdot T^{2H} > 0. \qquad [1.14]$$

PROOF.– Note that our kernel $z(t,s)$ is homogeneous in the following sense:

$$z(t,s) = T^{H-1/2} z(t/T, s/T).$$

Therefore,

$$\sup_{0\le t\le T} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2 = \sup_{0\le t\le T} \mathrm{E}\int_0^t (z(t,s)-a(s))^2 ds$$

$$= T^{2H-1} \sup_{0\le t\le T} \mathrm{E}\int_0^{t/T\cdot T} (z(t/T,s/T) - a(s/T\cdot T)T^{1/2-H})^2 ds$$

$$= T^{2H} \sup_{0\le u\le 1} \mathrm{E}\int_0^u (z(u,v) - a(v\cdot T)T^{1/2-H})^2 dv,$$

and consequently ρ_T can be rewritten via ρ_1, namely $\rho_T = T^{2H}\rho_1$. This implies that we can restrict the consideration to the case of ρ_T with $T=1$. Now we construct a lower bound for

$$\max_{0\le t\le 1} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2 = \max_{0\le t\le 1} \int_0^t \big(z(t,s)-a(s)\big)^2\, ds.$$

Let $0 < t_1 \le 1$. Consider the random variable $\int_0^1 a(s)dW_s =: B$. Then,

$$\max_{0\le t\le 1} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2 = \max_{0\le t\le 1} \mathrm{E}\left(B_t^H - \mathrm{E}[B \mid \mathcal{F}_t]\right)^2$$
$$\ge \max\left\{\mathrm{E}\left(B_{t_1}^H - \mathrm{E}[B \mid \mathcal{F}_{t_1}]\right)^2, \mathrm{E}\left(B_1^H - B\right)^2\right\}.$$

Now we can use variance partitioning. Obviously, for every square integrable random variable η, we have

$$\mathrm{E}\left[(\eta - \mathrm{E}[\eta \mid B_{t_1}])^2 \mid B_{t_1}\right] = \mathrm{E}\left[\eta^2 \mid B_{t_1}\right] - (\mathrm{E}[\eta \mid B_{t_1}])^2.$$

Hence,

$$\mathrm{E}(\eta - \mathrm{E}[\eta \mid B_{t_1}])^2 = \mathrm{E}\eta^2 - \mathrm{E}(\mathrm{E}[\eta \mid B_{t_1}])^2.$$

We apply the inequality $\mathrm{E}\eta^2 \ge \mathrm{E}(\mathrm{E}[\eta \mid B_{t_1}])^2$ for $\eta = B_{t_1}^H - \mathrm{E}[B \mid \mathcal{F}_{t_1}]$ and for $\eta = B_1^H - B$, and obtain

$$\max_{0\le t\le 1} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2$$
$$\ge \max\left\{\mathrm{E}\left(B_{t_1}^H - \mathrm{E}\left[B \mid B_{t_1}^H\right]\right)^2, \mathrm{E}\left(\mathrm{E}\left[B_1^H \mid B_{t_1}^H\right] - \mathrm{E}\left[B \mid B_{t_1}^H\right]\right)^2\right\}$$
$$\ge \frac{1}{2}\left(\mathrm{E}\left(B_{t_1}^H - \mathrm{E}\left[B \mid B_{t_1}^H\right]\right)^2 + \mathrm{E}\left(\mathrm{E}\left[B_1^H \mid B_{t_1}^H\right] - \mathrm{E}\left[B \mid B_{t_1}^H\right]\right)^2\right).$$

Note that for all real numbers P, Q and r, the inequality

$$\frac{(P-r)^2}{2} + \frac{(Q-r)^2}{2} \ge \frac{(P-Q)^2}{4} \qquad [1.15]$$

holds true because $2(P-r)^2 + 2(Q-r)^2 - (P-Q)^2 = (P+Q-2r)^2 \ge 0$. Therefore,

$$\max_{0\le t\le 1} \mathrm{E}\left(B_t^H - \int_0^t a(s)\,dW_s\right)^2$$
$$\ge \frac{1}{4}\mathrm{E}\left(B_{t_1}^H - \mathrm{E}\left[B_1^H \mid B_{t_1}^H\right]\right)^2 = \frac{1}{4}\mathrm{E}\left(B_{t_1}^H - \frac{\mathrm{E}\left(B_1^H B_{t_1}^H\right)}{\mathrm{E}\left(B_{t_1}^H\right)^2} B_{t_1}^H\right)^2$$

$$= \frac{1}{4}\mathrm{E}\left(B^H_{t_1}\left(1 - \frac{1 + t_1^{2H} - (1-t_1)^{2H}}{2t_1^{2H}}\right)\right)^2$$

$$= \frac{1}{4}t_1^{2H}\left(1 - \frac{1 + t_1^{2H} - (1-t_1)^{2H}}{2t_1^{2H}}\right)^2 = \frac{\left(1 - t_1^{2H} - (1-t_1)^{2H}\right)^2}{16t_1^{2H}}. \qquad \square$$

REMARK 1.1.– By substituting $t = \frac{1}{2}$ into maximized expression in [1.14], we obtain a loose lower bound

$$\rho_T \geq \frac{\left(2^{2H} - 2\right)^2}{16 \cdot 2^{2H}} \cdot T^{2H}.$$

1.4. The existence of minimizing function for the principal functional

Recall that in the analytic form, our goal is to find

$$\inf_{x \in L_2([0,T])} \sup_{t \in [0,T]} \int_0^t \left(z(t,s) - x(s)\right)^2 ds = \inf_{x \in L_2([0,T])} f(x),$$

where the functional $f = f(x)$ is defined via [1.1], and a minimizing element $x \in L_2([0,T])$ if the infimum is reached. In the original formulation, z is the kernel related to an fBm. However, the solution of this problem is based on the general properties of functionals in a Hilbert space, in particular, functionals defined by kernels satisfying the assumption **(A)** with inequality [1.2]. Therefore, in this section, we consider arbitrary kernel z satisfying assumption **(A)**, which implies that the functional $f = f(x)$ is well defined for any $x \in L_2([0,1])$. From this point on, with a view to simplifying the computations, let us consider in this chapter only the case $T = 1$.

LEMMA 1.2.– *For any $x, y \in L_2([0,1])$,*

$$|f(x) - f(y)| \leq \|x - y\|_{L_2([0,1])}. \qquad [1.16]$$

PROOF.– Evidently, for any $x, y \in L_2([0,1])$ and $0 \leq t \leq 1$,

$$\left(\int_0^t \left(z(t,s) - x(s)\right)^2 ds\right)^{1/2} \leq \left(\int_0^t \left(x(s) - y(s)\right)^2 ds\right)^{1/2} + \left(\int_0^t \left(z(t,s) - y(s)\right)^2 ds\right)^{1/2}.$$

Therefore,

$$\sup_{t\in[0,1]}\left(\int_0^t \big(z(t,s)-x(s)\big)^2\,ds\right)^{1/2}$$
$$\le \sup_{t\in[0,1]}\left(\int_0^t \big(x(s)-y(s)\big)^2\,ds\right)^{1/2}+\sup_{t\in[0,1]}\left(\int_0^t \big(z(t,s)-y(s)\big)^2\,ds\right)^{1/2},$$

which is clearly equivalent to the inequality

$$f(x)\le \|x-y\|_{L_2([0,1])}+f(y).$$

Swapping x and y, we establish [1.16] and thus obtain the proof. □

COROLLARY 1.2.– *The functional f is continuous on $L_2([0,1])$.*

LEMMA 1.3.– *The following inequalities hold for any function $x\in L_2([0,1])$:*

$$\big|\,\|x\|_{L_2([0,1])}-\|z(1,\cdot)\|_{L_2([0,1])}\,\big|\le f(x)\le \|x\|_{L_2([0,1])}+f(0). \qquad [1.17]$$

PROOF.– The left-hand side of [1.17] immediately follows from the inequalities

$$f(x)\ge\left(\int_0^1 \big(z(1,s)-x(s)\big)^2\,ds\right)^{1/2}=\|z(1,\cdot)-x\|_{L_2([0,1])}$$
$$\ge\big|\,\|x\|_{L_2([0,1])}-\|z(1,\cdot)\|_{L_2([0,1])}\,\big|,$$

and the right-hand side of [1.17] follows from [1.16]. □

LEMMA 1.4.– *The functional f is convex on $L_2([0,1])$.*

PROOF.– We have to prove that for any $x,y\in L_2([0,1])$ and any $\alpha\in[0,1]$,

$$f\big(\alpha x+(1-\alpha)y\big)\le \alpha f(x)+(1-\alpha)f(y). \qquad [1.18]$$

Applying the triangle inequality, we note that for any $t\in[0,1]$

$$\left(\int_0^t \big[\alpha x(s)+(1-\alpha)y(s)-z(t,s)\big]^2\,ds\right)^{\frac12}$$
$$\le\left(\int_0^t \big[\alpha\left(z(t,s)-x(s)\right)\big]^2\,ds\right)^{\frac12}+\left(\int_0^t \big[(1-\alpha)\big(z(t,s)-y(s)\big)\big]^2\,ds\right)^{\frac12},$$

whence

$$\sup_{t\in[0,1]}\left(\int_0^t \left[\alpha x(s)+(1-\alpha)y(s)-z(t,s)\right]^2 ds\right)^{\frac{1}{2}}$$

$$\leq \alpha \sup_{t\in[0,1]}\left(\int_0^t \left(z(t,s)-x(s)\right)^2 ds\right)^{\frac{1}{2}}$$

$$+ (1-\alpha) \sup_{t\in[0,1]}\left(\int_0^t \left(z(t,s)-y(s)\right)^2 ds\right)^{\frac{1}{2}},$$

and inequality [1.18] follows. □

THEOREM 1.3.– *The functional f reaches its minimal value on $L_2([0,1])$.*

PROOF.– By Corollary 1.2 and Lemma 1.4 the functional f is continuous and convex. By Lemma 1.3, $f(x)$ tends to $+\infty$ as $\|x\| \to \infty$. Hence, it follows from Proposition A1.2 that f reaches its minimal value. □

1.5. An example of the principal functional with infinite set of minimizing functions

We continue to study arbitrary kernel z satisfying assumption **(A)**, which implies that the functional f is well defined for any $x \in L_2([0,1])$. Note that the set $\mathfrak{M}_f$ of minimizing functions for the functional f is convex. In this section, we consider an example of kernel z for which $\mathfrak{M}_f$ contains more than one point and consequently is infinite. First, we establish the following lower bound for the functional f, which is similar to the particular case, considered in Theorem 1.2.

LEMMA 1.5.– *1) Let the kernel z of the functional f defined by* [1.1] *satisfy assumption* **(A)**. *Then for any $a \in L_2([0,1])$ and $0 \leq t_1 < t_2 \leq 1$, the following inequality holds*

$$\sup_{t\in[0,1]}\int_0^t \left(z(t,s)-a(s)\right)^2 ds \geq \frac{1}{4}\int_0^{t_1}\left(z(t_2,s)-z(t_1,s)\right)^2 ds. \qquad [1.19]$$

2) The equality in [1.19] *implies that*

$$a(s) = \frac{1}{2}\left(z(t_1,s)+z(t_2,s)\right) \quad a.e. \text{ on } [0,t_1), \qquad [1.20]$$

and

$$a(s) = z(t_2,s) \quad a.e. \text{ on } [t_1,t_2]. \qquad [1.21]$$

PROOF.– 1) The following inequalities are evident:

$$\begin{aligned}
&\sup_{t\in[0,1]} \int_0^t \big(z(t,s)-a(s)\big)^2\,ds \\
&\quad\geq \max\left\{\int_0^{t_1} \big(z(t_1,s)-a(s)\big)^2\,ds,\ \int_0^{t_2} \big(z(t_2,s)-a(s)\big)^2\,ds\right\} \\
&\quad\geq \max\left\{\int_0^{t_1} \big(z(t_1,s)-a(s)\big)^2\,ds,\ \int_0^{t_1} \big(z(t_2,s)-a(s)\big)^2\,ds\right\} \\
&\quad\geq \frac{1}{2}\int_0^{t_1} \Big(\big(z(t_1,s)-a(s)\big)^2+\big(z(t_2,s)-a(s)\big)^2\Big)\,ds. \qquad [1.22]
\end{aligned}$$

Setting in the inequality [1.15] $P = z(t_1, s)$, $Q = z(t_2, s)$ and $r = a(s)$, we obtain from [1.22] that

$$\begin{aligned}
&\sup_{t\in[0,1]} \int_0^t \big(z(t,s)-a(s)\big)^2 ds \\
&\quad\geq \frac{1}{2}\int_0^{t_1} \Big[\big(z(t_1,s)-a(s)\big)^2+\big(z(t_2,s)-a(s)\big)^2\Big]\,ds \\
&\quad\geq \frac{1}{4}\int_0^{t_1} \big(z(t_2,s)-z(t_1,s)\big)^2 ds. \qquad [1.23]
\end{aligned}$$

Thus, inequality [1.19] is proved.

2) Now we show that equality in [1.19] implies [1.20] and [1.21]. Indeed, equality in [1.15] holds if and only if $P + Q - 2r = 0$. Equality in [1.23] has a form

$$\begin{aligned}
&\frac{1}{2}\int_0^{t_1} \Big[\big(z(t_1,s)-a(s)\big)^2+\big(z(t_2,s)-a(s)\big)^2\Big]\,ds \\
&\quad= \frac{1}{4}\int_0^{t_1} \big(z(t_1,s)-z(t_2,s)\big)^2 ds,
\end{aligned}$$

and it holds if and only if

$$z(t_1,s)+z(t_2,s)-2a(s)=0 \quad \text{a.e. on } [0,t_1),$$

i.e. it holds if and only if condition [1.20] holds.

If [1.20] holds, then

$$\int_0^{t_1} \big(z(t_1,s)-a(s)\big)^2 ds = \frac{1}{4}\int_0^{t_1} \big(z(t_1,s)-z(t_2,s)\big)^2\, ds,$$

and

$$\int_0^{t_2} \big(z(t_2,s)-a(s)\big)^2 ds = \frac{1}{4}\int_0^{t_1} \big(z(t_2,s)-z(t_1,s)\big)^2 ds + \int_{t_1}^{t_2} \big(z(t_2,s)-a(s)\big)^2 ds.$$

It means that under condition [1.20], equality [1.19] holds if and only if

$$\int_{t_1}^{t_2} (z(t_2,s)-a(s))^2 ds = 0,$$

i.e. if and only if [1.21] holds. □

REMARK 1.2.– Let the kernel z of the functional f from [1.1] satisfy assumption **(A)**. Then, for any $a \in L_2([0,1])$ and $0 \le t_1 < t_2 \le 1$

$$\max_{t\in\{t_1,t_2\}} \int_0^t \big(z(t,s)-a(s)\big)^2\, ds \ge \frac{1}{4}\int_0^{t_1} \big(z(t_2,s)-z(t_1,s)\big)^2\, ds. \qquad [1.24]$$

Equality in [1.24] holds if and only if [1.20] and [1.21] hold.

THEOREM 1.4.– *(Example of functional f with infinite set $\mathfrak{M}_f$.) Take the kernel $z(t,s)$ of the form $z(t,s) = g(t)h(s)$, $t,s \in [0,1]$, where*

$$g(t) = (6t-2)\mathbf{1}_{\frac{1}{3}\le t\le\frac{1}{2}} + (4-6t)\mathbf{1}_{\frac{1}{2}\le t\le\frac{5}{6}} + (6t-6)\mathbf{1}_{\frac{5}{6}\le t\le 1}$$

and

$$h(s) = 4s\mathbf{1}_{0\le s\le\frac{1}{4}} + (2-4s)\mathbf{1}_{\frac{1}{4}\le s\le\frac{1}{2}}$$

(see Figure 1.1). Then,

$$\min_{a\in L_2([0,1])} \max_{t\in[0,1]} \int_0^t \big(z(t,s)-a(s)\big)^2 ds = \frac{1}{6}, \qquad [1.25]$$

and $\mathfrak{M}_f$ consists of functions $a(s)$ satisfying the conditions

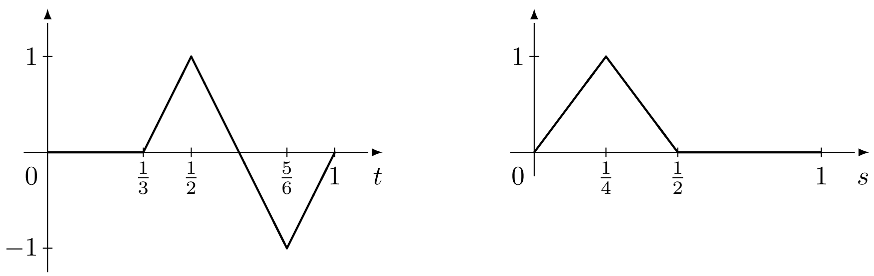
$$a(s) = 0 \quad a.e.\ on\ \left[0,\tfrac{5}{6}\right] \qquad [1.26]$$

and

$$\int_{\frac{5}{6}}^{t} a(s)^2 ds \le \tfrac{1}{6} - 6(1-t)^2, \quad \tfrac{5}{6} \le t \le 1. \qquad [1.27]$$

REMARK 1.3.– 1) Since $z \in C([0,1]^2)$ and $a \in L_2([0,1])$, we have that $\int_0^t \big(z(t,s) - a(s)\big)^2 ds$ is continuous in t. Therefore, we can really replace $\sup_{t\in[0,1]}$ with $\max_{t\in[0,1]}$ in equality [1.25].

2) Some examples of functions satisfying [1.26] and [1.27]: $a(s) = 0, s \in [0,1]$; $a(s) = (12(1-s))^{1/2}\mathbf{1}_{5/6<s\le 1}$; $a(s) = \sqrt{3}(6s-5)\mathbf{1}_{5/6\le s\le 1}$.



(a) Graph of function g.

(b) Graph of function h.

Figure 1.1. *Graphs of functions g and h.*

PROOF.– To establish a lower bound of the left-hand side of [1.25], note that

$$\int_0^t h(s)^2 ds = \frac{1}{6} \quad \text{for} \quad \frac{1}{2} \le t \le 1.$$

Therefore, applying Lemma 1.5 with $t_1 = \frac{1}{2}$ and $t_2 = \frac{5}{6}$ we obtain that

$$\sup_{t\in[0,1]} \int_0^t \big(z(t,s) - a(s)\big)^2 ds \ge \frac{1}{4}\int_0^{\frac{1}{2}} \big(z(5/6,s) - z(1/2,s)\big)^2 ds$$

$$= \frac{1}{4}\int_0^{\frac{1}{2}} \big(g(5/6)h(s) - g(1/2)h(s)\big)^2 ds = \frac{1}{4}\int_0^{\frac{1}{2}} 4h(s)^2 ds = \frac{1}{6}. \qquad [1.28]$$

Let us prove that functions $a = a(s)$ satisfying [1.26] and [1.27] transform [1.28] into equality.

To establish an upper bound of the left-hand side of [1.25], consider functions satisfying conditions [1.26] and [1.27]. Then, for $0 \le t \le \frac{5}{6}$ we have that

$$\int_0^t \big(z(t,s) - a(s)\big)^2 \, ds = \int_0^t z(t,s)^2 \, ds = \int_0^t g(t)^2 h(s)^2 \, ds$$
$$= g(t)^2 \int_0^t h(s)^2 \, ds \le \int_0^{\frac{5}{6}} h(s)^2 ds = \frac{1}{6},$$

since $a(s) = 0$ on $\left[0, \frac{5}{6}\right]$ and $g(t)^2 \le 1$. For $\frac{5}{6} < t \le 1$, we take into account the values of a, h and g on this interval and obtain that

$$\begin{aligned}\int_0^t \big(z(t,s) - a(s)\big)^2 \, ds &= \int_0^{\frac{5}{6}} \big(g(t)h(s) - a(s)\big)^2 \, ds \\ &\quad + \int_{\frac{5}{6}}^t \big(g(t)h(s) - a(s)\big)^2 \, ds \\ &= \int_0^{\frac{5}{6}} g(t)^2 h(s)^2 \, ds + \int_{\frac{5}{6}}^t a(s)^2 \, ds \\ &\le (6t-6)^2 \cdot \frac{1}{6} + \frac{1}{6} - 6(1-t)^2 = \frac{1}{6}.\end{aligned} \qquad [1.29]$$

Hence, if the function a satisfies [1.26] and [1.27], we have that

$$\sup_{t \in [0,1]} \int_0^t \big(z(t,s) - a(s)\big)^2 \, ds \le \frac{1}{6}.$$

Summing up, we obtain [1.25].

Now we prove that any minimizing function a satisfies [1.26] and [1.27]. Indeed, let

$$\sup_{t \in [0,1]} \int_0^t \big(z(t,s) - a(s)\big)^2 \, ds = \frac{1}{6}.$$

Then inequality [1.28] is transformed into equality; therefore,

$$\sup_{t \in [0,1]} \int_0^t \big(z(t,s) - a(s)\big)^2 \, ds = \frac{1}{4} \int_0^{\frac{1}{2}} \Big(z\left(\tfrac{5}{6}, s\right) - z\left(\tfrac{1}{2}, s\right)\Big)^2 ds. \qquad [1.30]$$

It follows from [1.30] and from the second part of Lemma 1.5 that

$$a(s) = \frac{1}{2}\big(z(5/6, s) + z(1/2, s)\big) = \frac{1}{2}\big(g(5/6) + g(1/2)\big)h(s) = 0$$

a.e. on $\left[0, \frac{1}{2}\right]$ because $g(1/2) = 1$, $g(5/6) = -1$; we also obtain the equality

$$a(s) = z(5/6, s) = g(5/6)h(s) = 0$$

a.e. on $\left[\frac{1}{2}, \frac{5}{6}\right]$ because $h(s) = 0$ for $s \geq \frac{1}{2}$. Therefore, function a satisfies condition [1.26]. Then, similarly to the calculations provided in [1.29], we obtain

$$\int_0^t \big(z(t, s) - a(s)\big)^2 \, ds = \frac{(6t-6)^2}{6} + \int_{\frac{5}{6}}^t a(s)^2 \, ds \quad \text{for} \quad \frac{5}{6} < t \leq 1,$$

and it follows from the inequality $\int_0^t \big(z(t, s) - a(s)\big)^2 \, ds \leq \frac{1}{6}$ that

$$\int_{\frac{5}{6}}^t a(s)^2 \, ds \leq \frac{1}{6} - \frac{(6t-6)^2}{6} = \frac{1}{6} - 6(1-t)^2 \quad \text{for} \quad \frac{5}{6} < t \leq 1.$$

It means that the function a satisfies condition [1.27]. □

1.6. Uniqueness of the minimizing function for functional with the Molchan kernel and $H \in \left(\frac{1}{2}, 1\right)$

Now we return to the problem [1.12] of the approximation of an fBm by martingales.

First, we prove some simple but useful properties of the fractional Brownian kernel z (Molchan kernel), defined by [1.7].

LEMMA 1.6.– *(Molchan kernel properties)*

1) The kernel z satisfies condition **(A)**.

2) The kernel z increases in the first argument and decreases in the second argument.

3) The kernel $z \in C([0, 1]^2)$.

4) For any $c > 0$ and $0 < s \leq t$, $z(ct, cs) = c^\alpha z(t, s)$ with $\alpha = H - 1/2$.

PROOF.– 1. Since z is the Molchan kernel of an fBm, we have that

$$t^{2H} = \mathrm{E}\left(B_t^H\right)^2 = \mathrm{E}\left(\int_0^t z(t,s)\,dW_s\right)^2 = \int_0^t z(t,s)^2\,ds.$$

Therefore, $\sup_{t\in[0,1]}\int_0^t z(t,s)^2\,ds = 1$, and [1.2] is satisfied. Other statements follow directly from [1.7]. □

Now we are in position to establish the uniqueness of minimizing function for the principal functional corresponding to the Molchan kernel of fBm. In order to do this, first, we prove an auxiliary statement concerning any minimizing function for this functional. For $x \in L_2([0,1])$, denote

$$g_x(t) = \left(\int_0^t \left(z(t,s) - x(s)\right)^2 ds\right)^{1/2}.$$

Then, we have from the definition of the principal functional f that $f(x) = \sup_{t\in[0,1]} g_x(t)$. It follows from Lemma 1.6 that $g_x \in C[0,1]$ for any $x \in L_2([0,1])$. Using the self-similarity property 4) of the kernel z, it is easy to note that

$$g_a(t) = c^{\alpha+1/2} g_{c^{-\alpha}a(c\cdot)}(t/c). \qquad [1.31]$$

LEMMA 1.7.– *Let $a \in \mathfrak{M}_f$. Then, the maximal value of g_a is reached at point 1, i.e. $f(a) = g_a(1)$.*

PROOF.– Set $a(t) = 0$ for $t > 1$. Suppose that $g_a(1) < f(a)$. Since $g_a(t)$ is continuous in t, $c > 1$ exists such that $g_a(t) < f_a$ for $t \in [1,c]$. This means that $\max_{t\in[0,c]} g_a(t) = f_a$. Set $b(t) = c^{-\alpha}a(tc)$. It follows from equation [1.31] that $g_b(t) = c^{-1/2-\alpha} g_a(tc)$, $t \in [0,1]$. We immediately obtain $f(b) = c^{-\alpha-1/2}f(a) < f(a)$, which leads to a contradiction. □

REMARK 1.4.– Similarly, if the function g_a is differentiable at point 1, then $g_a'(1) \geq 2Hg_a(1)$.

THEOREM 1.5.– *(Uniqueness of a minimizing function) For the principal functional f defined by* [1.1] *with the fractional Brownian kernel z from* [1.7]*, there is a unique minimizing function.*

PROOF.– Let us denote M_f as the minimal value of functional f. Recall that the set $\mathfrak{M}_f$ is non-empty and convex. Let $\hat{z}(s) = z(1,s)$, $s \in [0,1]$. It follows from Lemma 1.7 that for any function $x \in \mathfrak{M}_f$ the following equality holds:

$$f(x) = \left(\int_0^1 \left(x(s) - z(1,s)\right)^2 ds\right)^{1/2} = \|x - \hat{z}\|_{L_2([0,1])}.$$

For any $x, y \in \mathfrak{M}_f$, $\alpha \in (0,1)$, we have that $\alpha x + (1-\alpha)y \in \mathfrak{M}_f$. Hence,

$$\begin{aligned} M_f &= f\big(\alpha x + (1-\alpha)y\big) = \|\alpha x + (1-\alpha)y - \hat{z}\|_{L_2([0,1])} \\ &\le \alpha \|x - \hat{z}\|_{L_2([0,1])} + (1-\alpha)\|y - \hat{z}\|_{L_2([0,1])} \\ &= \alpha f(x) + (1-\alpha) f(y) = M_f. \end{aligned}$$

Thus,

$$\|\alpha x + (1-\alpha)y - \hat{z}\|_{L_2([0,1])} = \alpha \|x - \hat{z}\|_{L_2([0,1])} + (1-\alpha)\|y - \hat{z}\|_{L_2([0,1])}$$

for all $\alpha \in (0,1)$, in particular,

$$\begin{aligned} &\left\|\tfrac{1}{2}x + \tfrac{1}{2}y - \hat{z}\right\|_{L_2([0,1])} = \tfrac{1}{2}\|x - \hat{z}\|_{L_2([0,1])} + \tfrac{1}{2}\|y - \hat{z}\|_{L_2([0,1])}, \\ &\|x - \hat{z} + y - \hat{z}\|_{L_2([0,1])} = \|x - \hat{z}\|_{L_2([0,1])} + \|y - \hat{z}\|_{L_2([0,1])}. \end{aligned}$$

For arbitrary vectors x and y in a Hilbert space, the equality $\|x + y\| = \|x\| + \|y\|$ implies that x and y are parallel and in the same direction (or either one or both of them are zero vectors). Therefore, the functions $x - \hat{z}$ and $y - \hat{z}$ are equal up to a non-negative multiplier, $x - \hat{z} = C(y - \hat{z})$ or $C(x - \hat{z}) = y - \hat{z}$, with $C \ge 0$, but since $\|x - \hat{z}\|_{L_2([0,1])} = \|y - \hat{z}\|_{L_2([0,1])}$, we have $x - \hat{z} = y - \hat{z}$. Therefore, $x = y$ (in $L_2([0,1])$), i.e. a.e. on $[0,1]$, as required. □

REMARK 1.5.– In this section, we assume that $\frac{1}{2} < H < 1$. However, Theorem 1.5 is valid for all H, $0 < H < 1$. Indeed, statements 1 and 4 of Lemma 1.6 hold true for all $H \in (0,1)$. In the proof, we refer to equation [1.5] instead of [1.7]. Statements 2) and 3) of Lemma 1.6 do not hold true for $0 < H < \frac{1}{2}$. Lemma 1.7 also holds true for all $H \in (0,1)$, and Theorem 1.5 follows from Lemma 1.7.

THEOREM 1.6.– *Let a be the function that minimizes f, i.e. $a \in \mathfrak{M}_f$. Then, a function $\phi\colon [0,1] \to \mathbb{R}$ exists such that $s \le \phi(s) \le 1$, $s \in [0,1]$, and $a(s) = z(\phi(s), s)$ a.e.*

PROOF.– Since $z(t,s) \ge 0$ for all t and s, we have that

$$(z(t,s) - a(s))^2 \ge (z(t,s) - \max(0, a(s)))^2, \qquad t, s \in [0,1],$$

whence

$$\int_0^t (z(t,s) - a(s))^2\, ds \ge \int_0^t (z(t,s) - \max(0, a(s)))^2\, ds, \qquad t \in [0,1].$$

Finally,

$$f(a) \geq f(\max(a, 0)).$$

It means that function $\max(a, 0) \in \mathfrak{M}_f$. Due to the uniqueness of the minimizing function (see Theorem 1.5), $a(s) = \max(a(s), 0)$ for almost all (a.a.) s, whence $a(s) \geq 0$ a.e. on $[0, 1]$.

Furthermore, since the kernel $z(t, s)$ increases in t, we have

$$z(t, s) \leq z(1, s), \qquad t, s \in [0, 1],$$

whence

$$(z(t, s) - a(s))^2 \geq (z(t, s) - \min(z(1, s), a(s)))^2, \qquad t, s \in [0, 1].$$

It means that

$$\int_0^t (z(t, s) - a(s))^2 \, ds \geq \int_0^t (z(t, s) - \min(z(1, s), a(s)))^2 \, ds, \qquad t \in [0, 1],$$

and finally,

$$f(a) \geq f(\min(\hat{z}, a)),$$

where $\hat{z}(s) = z(1, s)$. Thus, function $\min(\hat{z}, a) \in \mathfrak{M}_f$. Due to the uniqueness of the minimizing function, $a(s) = \min(z(1, s), a(s))$ for a.a. s, whence $a(s) \leq z(1, s)$ a.e. on $[0, 1]$.

We have just proved that

$$0 = z(s, s) \leq a(s) \leq z(1, s) \quad \text{a.e. in } [0, 1].$$

Since the kernel $z(t, s)$ is continuous in t, there exists a function $\phi(s)$, $s \leq \phi(s) \leq 1$, such that $a(s) = z(\phi(s), s)$ for a.a. $s \in [0, 1]$. $\square$

COROLLARY 1.3.– *The minimizing function $a \in \mathfrak{M}_f$ is non-negative. It is fully consistent with Corollary 1.1, where it is stated that we can restrict ourselves to non-negative functions a.*

1.7. Representation of the minimizing function

The main problem of our minimization procedure is that the minimizing function has no explicit analytical representation. Therefore, we can only study its properties and give the approximation formulae. In this section we consider principal functional f corresponding to fBm and establish that the minimizing function has a special form. We start by proving several auxiliary results concerning the Molchan kernel and the minimizing function.

1.7.1. *Auxiliary results*

LEMMA 1.8.– *For any $0 \le t \le 1$ Molchan kernel satisfies the relation*

$$\int_0^t \big(z(1,s) - z(t,s)\big)^2 \, ds + \int_t^1 z(1,s)^2 \, ds = (1-t)^{2H}. \qquad [1.32]$$

PROOF.– By [1.4], the left-hand side of [1.32] is equal to $\mathrm{E}\left(B_1^H - B_t^H\right)^2 = (1-t)^{2H}$. □

The next statement will be essentially generalized in what follows. However, we prove it because its proof clarifies the main ideas; moreover; it has interesting consequences concerning the properties of the minimizing function. In the remainder of this section, $a = a(s)$, $s \in [0,1]$ denotes the minimizing function, i.e. the unique element of $\mathfrak{M}_f$.

LEMMA 1.9.– *Let $t^* = \sup\{t \in (0,1) : g_a(t) = f(a)\}$ ($t^* = 0$ if this set is empty). If $t^* < 1$, then $a(t) = z(1,t)$ for a.e. $t \in [t^*, 1]$.*

PROOF.– Fix some $t_1 \in (t^*, 1]$ and prove that for any $h \in L_2([0,1])$ the following equality holds:

$$\int_{t_1}^1 h(s)\big(a(s) - z(1,s)\big) \, ds = 0.$$

Evidently, the proof follows immediately from this statement.

Assume the contrary. Then, without loss of generality, $h \in L_2([0,1])$ exists such that

$$\int_{t_1}^1 h(s)\big(a(s) - z(1,s)\big) \, ds =: \kappa > 0.$$

It follows from the continuity of the last integral w.r.t. upper bound that for some $t_2 \in (t_1, 1]$ we have

$$\int_{t_1}^{t} h(s)\big(a(s) - z(t,s)\big)\, ds \geq \kappa/2$$

for any $t \in [t_2, 1]$. Note also that our assumption implies that

$$m := \max_{s\in[t_1,t_2]} g_a(s) < f(a).$$

Consider now $b_\delta(t) = a(t) - \delta h(t)\mathbf{1}_{[t_1,1]}(t)$ for $\delta > 0$. We have that $g_{b_\delta}(t) = g_a(t)$ for $t \in [0, t_1]$, and

$$g_{b_\delta}(t)^2 = g_a(t)^2 - 2\delta \int_{t_1}^{t} h(s)\big(a(s) - z(t,s)\big)\, ds + \delta^2 \int_{t_1}^{t} h(s)^2\, ds$$

for $t > t_1$. For $t \in (t_1, t_2]$ the following inequality holds:

$$g_{b_\delta}(t)^2 \leq m^2 - 2\delta \int_{t_1}^{t} h(s)\big(a(s) - z(t,s)\big)\, ds + \delta^2 \int_{t_1}^{t} h(s)^2\, ds \leq m^2 + C\delta$$

with the constant C that does not depend on t, δ. Then, for sufficiently small $\delta > 0$, we have that $g_{b_\delta}(t) < f(a)$ for any $t \in (t_1, t_2]$.

Furthermore, if $t \in (t_2, 1]$, then

$$\begin{aligned} g_{b_\delta}(t)^2 &\leq f(a)^2 - 2\delta \int_{t_1}^{t} h(s)\big(a(s) - z(t,s)\big)\, ds + \delta^2 \int_{t_1}^{t} h(s)^2\, ds \\ &\leq f(a)^2 - \kappa\delta + \delta^2 \int_0^1 h(s)^2\, ds. \end{aligned}$$

Again, for sufficiently small $\delta > 0$ and any $t \in (t_2, 1]$, we have that $g_{b_\delta}(t) < f(a)$. Therefore, for sufficiently small $\delta > 0$, we obtain $f(b_\delta) = f(a)$ and $g_{b_\delta}(1) < f(a) = f(b_\delta)$. We obtain the contradiction with Lemma 1.7 whence the proof follows. □

COROLLARY 1.4.– *A point $t \in (0, 1)$ exists such that $g_a(t) = f(a)$.*

PROOF.– Assuming the contrary, we obtain from Lemma 1.9 that $a(t) = z(1, t)$ for a.a. $t \in [0, 1]$. However, in this case, $g_a(1) = 0$, which contradicts Lemma 1.7. □

Denote $\mathfrak{G}_a = \{t \in [0,1] : g_a(t) = f(a)\}$, the set of the maximal points of the function g_a.

LEMMA 1.10.– *Let $u \in [0,1)$ be a point such that $g_a(u) < f(a)$. Then, there does not exist a function $h \in L_2([0,1])$ such that for any $t \in \mathfrak{G}_a \cap (u,1]$ the inequality $\int_u^t h(s)\big(a(s) - z(t,s)\big)\, ds > 0$ holds.*

PROOF.– Assume the contrary, i.e. let for some function $h \in L_2([0,1])$ we have that $\int_u^t h(s)\big(a(s) - z(t,s)\big)ds > 0$ for any $t \in \mathfrak{G}_a \cap (u,1]$. The set $\mathfrak{G}_a \cap (u,1]$ is closed because $g_a(u) < f(a)$. Therefore,

$$\kappa := \min_{t \in \mathfrak{G}_a \cap (u,1]} \int_u^t h(s)\big(a(s) - z(t,s)\big)\, ds > 0.$$

Denote

$$\mathfrak{B}_\varepsilon = \{t \in (u,1] : \mathfrak{G}_a \cap (u,1] \cap (t-\varepsilon, t+\varepsilon) \neq \varnothing\}$$

the intersection of ε-neighborhood of the set $\mathfrak{G}_a \cap (u,1]$ with interval $(u,1]$. Continuity argument implies that for some $\varepsilon > 0$ it holds that

$$\int_u^t h(s)\big(a(s) - z(t,s)\big)\, ds > \kappa/2$$

for any $t \in \mathfrak{B}_\varepsilon$.

Similarly to the proof of Lemma 1.9, denote $b_\delta(t) = a(t) - \delta h(t)\mathbf{1}_{(u,1]}(t)$ for any $\delta > 0$. Then, we have that $g_{b_\delta}(t) = g_a(t)$ for any $t \in [0,u]$, and

$$\begin{aligned} g_{b_\delta}(t)^2 &\le f(a)^2 - 2\delta \int_u^t h(s)\big(a(s) - z(t,s)\big)\, ds + \delta^2 \int_u^t h(s)^2\, ds \\ &\le f(a)^2 - \kappa\delta + \delta^2 \int_0^1 h(s)^2\, ds \end{aligned}$$

for any $t \in \mathfrak{B}_\varepsilon$. It follows from the continuity of g_a that

$$m = \max_{t \in [u,1] \setminus \mathfrak{B}_\varepsilon} g_a(t) < f(a).$$

Therefore, we have for $t \in (u,1] \setminus \mathfrak{B}_\varepsilon$ that

$$\begin{aligned} g_{b_\delta}(t)^2 &= g_a(t)^2 - 2\delta\big(a(s) - z(1,s)\big) + \delta^2 \int_{t_1}^t h(s)^2\,ds \\ &\le m^2 - 2\delta \int_{t_1}^t h(s)\big(a(s) - z(t,s)\big)\,ds + \delta^2 \int_{t_1}^t h(s)^2\,ds \\ &\le m^2 + C\delta, \end{aligned}$$

with the constant C that does not depend on t and δ. It follows from the above bounds that for sufficiently small $\delta > 0$ and for any $t \in (u,1]$ we have the inequality $g_{b_\delta}(t) < f(a)$. It means that for sufficiently small $\delta > 0$, we obtain the equality $f(b_\delta) = f(a)$, and moreover, $g_{b_\delta}(1) < f(a) = f(b_\delta)$, which contradicts Lemma 1.7. □

Lemma 1.10 supplies the form of minimizing function on the part of the interval $[0,1]$. All equalities below are considered almost surely (a.s).

LEMMA 1.11.– *Let $t_1 = \min\{t \in (0,1) : g_a(t) = f(a)\}$. Then, $t_2 \in (t_1,1] \cap \mathfrak{G}_a$ and a random variable ξ_a exist with the values in $[t_1,t_2] \cap \mathfrak{G}_a$ such that for $t \in [0,t_2)$, we have that $\mathrm{P}(\xi_a \ge t) > 0$, and the equality*

$$a(t) = \mathrm{E}\big[z(\xi_a,t) \bigm| \xi_a \ge t\big]$$

holds.

PROOF.– Consider the set of functions

$$\mathcal{K} = \{k_t(s) = z(t,s)\mathbf{1}_{s\le t} + a(s)\mathbf{1}_{s>t},\ t \in \mathfrak{G}_a\}$$

and let

$$\mathcal{C} = \left\{\int_0^1 k_t(s)\,F(dt),\ F \text{ is a distribution function on } \mathfrak{G}_a\right\}$$

be the closure of the convex hull of $\mathcal{K}$. According to Lemma 1.10, applied to $u = 0$, there does not exist $h \in L_2([0,1])$ such that $(h,k) < (h,a)$ for any $k \in \mathcal{K}$. Moreover, there is no $h \in L_2([0,1])$ such that $(h,k) < (h,a)$ for any $k \in \mathcal{C}$, i.e. the element a and the set $\mathcal{K}$ cannot be separated properly. Then, according to the proper separation theorem (see Corollary A1.2 in Appendix 1), $a \in \mathcal{C}$, so there exists a distribution F on $\mathfrak{G}_a$ such that

$$a(s) = \int_0^1 k_t(s)\,G(dt) = \int_{[s,1]} k_t(s)\,F(dt) + \int_{[0,s)} a(s)\,F(dt). \qquad [1.33]$$

Hence,

$$a(s)F([s,1]) = \int_{[s,1]} k_t(s)\, F(dt). \tag{1.34}$$

Note that the equality $\operatorname{supp} F = \{t_1\}$ is impossible because otherwise it follows from equation [1.34] that $a(s) = z(t_1, s)$ for $s \le t_1$; therefore, $g_a(t_1) = 0$ which contradicts the assumption $g_a(t) = f(a)$.

Using the latter statement and [1.34], we obtain the statement of the theorem with $t_2 = \max(\operatorname{supp} F)$ and a random variable ξ_a with the distribution F. □

Conditions on minimizing function from Lemma 1.11 are sufficient in the following sense.

LEMMA 1.12.– *Let $y \in L_2([0,1])$. Define the kernel $z_y(t,s)$ for $s,t \in [0,1]$ as*

$$z_y(t,s) = \begin{cases} z(t,s), & t \ge s, \\ y(s), & t < s. \end{cases}$$

Function y is the minimizing function of the principal functional f if and only if a random variable ξ_y exists, which takes values in $[0,1]$ such that the following conditions hold:

$$y(s) = \mathrm{E} z_y(\xi_y, s) \quad \text{for almost all } s \in [0,1], \tag{1.35}$$

$$g_y(\xi_y) = f(y) \quad a.s. \tag{1.36}$$

PROOF.– The necessity was proved in Lemma 1.11. Indeed, if $y = a$, let us take $\xi_y = \xi_a$, where ξ_a was obtained in the course of the proof of Lemma 1.11. Then condition [1.35] follows from equality [1.33], while condition [1.36] follows from the fact that $\xi_a \in \mathfrak{G}_a$.

The sufficiency is proved basically by reversing a proper separation argument from Lemma 1.11: if a function belongs to the convex set $\mathcal{C}$, then it cannot be properly separated from this set, which means that it is a minimizer. To make this idea rigorous, assume the contrary: let a function y satisfy [1.35] and [1.36], but $y \notin \mathfrak{M}_f$. Then, a function $a \in L_2([0,1])$ exists such that $f(y) > f(a)$ (e.g. we can take a as the minimizing function). Functional f^2 is convex, therefore

$$f\big(y + \delta(a-y)\big)^2 \le f(y)^2 + \delta\left(f(a)^2 - f(y)^2\right), \quad 0 \le \delta \le 1.$$

It is easy to see that for any function $b \in L_2([0,1])$,

$$\begin{aligned}
&\max_{t\in[0,1]} \|z_y(t,\cdot) - b\|^2 \\
&\quad = \max_{t\in[0,1]} \left(\int_0^t \big(z(t,s) - b(s)\big)^2 \, ds + \int_t^1 \big(y(s) - b(s)\big)^2 \, ds \right) \\
&\quad \le \max_{t\in[0,1]} \int_0^t \big(z(t,s) - b(s)\big)^2 \, ds + \int_0^1 \big(y(s) - b(s)\big)^2 \, ds \\
&\quad = f(b)^2 + \|y - b\|^2 .
\end{aligned}$$

Therefore, for $0 \le \delta \le 1$ we have that

$$\max_{t\in[0,1]} \|z_y(t,\cdot) - y - \delta(a-y)\|^2 \le f(y)^2 - \delta\left(f(y)^2 - f(a)^2\right) + \delta^2 \|a - y\|^2 .$$

It means that for sufficiently small $\delta > 0$

$$\max_{t\in[0,1]} \|z_y(t,\cdot) - y - \delta(a-y)\|^2 < f(y)^2. \qquad [1.37]$$

On the one hand, choose arbitrary δ for which inequality [1.37] holds, and set $b = y + \delta(a-y)$. Then,

$$\max_{t\in[0,1]} \|z_y(t,\cdot) - b\|^2 < f(y)^2. \qquad [1.38]$$

On the other hand,

$$\begin{aligned}
\max_{t\in[0,1]} \|z_y(t,\cdot) - b\|^2 &\ge \mathrm{E}\,\|z_y(\xi_y,\cdot) - b\|^2 \ge \mathrm{E}\,\|z_y(\xi_y,\cdot) - \mathrm{E}z_y(\xi_y,\cdot)\|^2 \\
&= \mathrm{E}\,\|z_y(\xi_y,\cdot) - y\|^2 = \mathrm{E}g_y(\xi_y)^2 = f(y)^2. \qquad [1.39]
\end{aligned}$$

Inequalities [1.38] and [1.39] contradict each other. Thus, assuming that the function y is not minimizing for the principal functional f, we obtain a contradiction. Therefore, $f(y) = \min f$. □

Now we are in position to prove that

$$\operatorname{ess\,sup} \xi_a := \min\{t : \mathrm{P}(\xi_a \le t) = 1\} = \max(\operatorname{supp} \xi_a) = 1,$$

which will imply that $t_2 = 1$ in Lemma 1.11.

LEMMA 1.13.– *Let a be the minimizing function for principal functional f and let ξ_a be a random variable satisfying conditions* [1.35] *and* [1.36] *where $y = a$. Then,* $\operatorname{ess\,sup} \xi_a = 1$.

PROOF.– Denote $t_2 = \operatorname{ess\,sup} \xi_a$. Evidently, ξ_a takes values from $[0, t_2]$.

Consider a function

$$b(s) = t_2^{-\alpha} a(t_2 s), \quad s \in [0,1].$$

Then, in view of the self-similarity property (item 4 in Lemma 1.6),

$$b(s) = \mathrm{E} z_b(\xi_a / t_2, s),$$

where $z_b(t,s)$ is defined in the formulation of Lemma 1.12. Using [1.31], we obtain

$$g_b(t) = t_2^{-H} g_a(t_2 t), \quad t \in [0,1].$$

On the one hand, since $a(s)$ satisfies [1.36], we have

$$f(b) = \max_{[0,1]} g_b \geq g_b\left(\xi_a / t_2\right) = t_2^{-H} g_a(\xi_a) = t_2^{-H} f(a)$$

a.s.; on the other hand,

$$f(b) = \max_{[0,1]} g_b = t_2^{-H} \max_{[0,t_2]} g_a \leq t_2^{-H} \max_{[0,1]} g_a = t_2^{-H} f(a).$$

This implies

$$f(b) = g_b\left(\xi_a / t_2\right) = t_2^{-H} f(a) \quad \text{a.e.}$$

Therefore, the function b satisfies [1.35] and [1.36] and is therefore a minimizer of f. Hence,

$$t_2^{-H} f(a) = f(b) = \min_{L_2([0,1])} f = f(a),$$

so $t_2 = 1$, as required. □

1.7.2. *Main properties of the minimizing function*

Now we can refine Lemma 1.11 in view of Lemma 1.13. Recall that a is the minimizing function for the principal functional f and $\mathfrak{S}_a = \{t \in [0,1] : g_a(t) = f(a)\}$.

THEOREM 1.7.– *A random variable $\widetilde{\xi}_a$ exists, taking values in $\mathfrak{S}_a$, such that*

$$\mathrm{P}(\widetilde{\xi}_a \geq s) > 0 \quad \textit{for all } s \in [0,1),$$
$$a(s) = \mathrm{E}[z(\widetilde{\xi}_a, s) \mid \widetilde{\xi}_a \geq s] \quad \text{a.e. } \textit{in } [0,1]. \qquad [1.40]$$

PROOF.– We can put $\tilde{\xi}_a = \xi_a$, where ξ_a comes from Lemma 1.11. Then the present statement is a straightforward consequence of Lemma 1.13. □

Without loss of generality we will assume in what follows that [1.40] holds for *every* $s \in [0,1]$:

$$a(s) = \mathrm{E}[z(\xi_a, s) \mid \xi_a \geq s] \quad \text{for any } s \in [0,1]. \qquad [1.41]$$

COROLLARY 1.5.– *1) The minimizing function a is left-continuous and has right limits.*

2) For any $s \in [0,1)$,

$$0 < a(s) \leq z(1,s), \qquad [1.42]$$

moreover,

$$a(s) < z(1,s)$$

on a set of positive Lebesgue measure.

PROOF.– 1) Follows from [1.41], continuity of z and the dominated convergence.

2) Taking into account statement 2 of Lemma 1.6, for $0 < s < t \leq 1$

$$0 < z(t,s) \leq z(1,s).$$

Now [1.42] follows from [1.41] and the fact that $\mathrm{P}(\xi_a > s) > 0$ for $s < 1$. Further, if $a(s) = z(1,s)$ a.e., then $g_a(1) = 0$, which contradicts Lemma 1.7. □

Further, we investigate the distribution of ξ.

LEMMA 1.14.– *There exists $t^* \in (0,1)$ such that for all $t \in (t^*,1)$, $g_a(t) < f(a)$.*

PROOF.– By Corollary 1.5, u_1 and u_2 exist such that $0 < u_1 < u_2 < 1$ and

$$\lambda_1 \{s \in [u_1, u_2] : a(s) < z(1,s)\} > 0,$$

where λ_1 is the Lebesgue measure.

Now estimate $f(a)^2 - g_a(t)^2$ from below, asymptotically as $t \to 1-$. For $u_2 \le t \le 1$ make simple transformations

$$\begin{aligned}
f(a)^2 - g_a(t)^2 &= g_a(1)^2 - g_a(t)^2 \\
&= \int_0^1 \big(z(1,s) - a(s)\big)^2 \, ds - \int_0^t \big(z(t,s) - a(s)\big)^2 \, ds \\
&= -\int_0^t \big(z(1,s) - z(t,s)\big)^2 \, ds + \int_t^1 \big(z(1,s) - a(s)\big)^2 \, ds \\
&\quad + 2\left(\int_0^{u_1} + \int_{u_2}^t\right) \big(z(1,s) - z(t,s)\big)\big(z(1,s) - a(s)\big) \, ds \\
&\quad + 2\int_{u_1}^{u_2} \big(z(1,s) - z(t,s)\big)\big(z(1,s) - a(s)\big) \, ds.
\end{aligned}$$

We estimate each of the terms separately. By Lemma 1.8,

$$\int_0^t \big(z(1,s) - z(t,s)\big)^2 ds < (1-t)^{2H}.$$

Furthermore,

$$\left(\int_0^{u_1} + \int_{u_2}^t\right) \big(z(1,s) - z(t,s)\big)\big(z(1,s) - a(s)\big) \, ds \ge 0,$$

since the integrand is non-negative by Lemmata 1.6 and 1.5;

$$\int_t^1 \big(z(1,s) - a(s)\big)^2 \ge 0.$$

It is easy to see that

$$\frac{z(1,s) - z(t,s)}{1-t} \to z_t'(1,s) = s^{-\alpha}(1-s)^{1-\alpha}$$

uniformly in $[u_1, u_2]$ as $t \to 1-$. Therefore,

$$\frac{\int_{u_1}^{u_2} \big(z(1,s) - z(t,s)\big)\big(z(1,s) - a(s)\big)\, ds}{1-t} \to \int_{u_1}^{u_2} z_t'(1,s)\big(z(1,s) - a(s)\big)\, ds,$$

as $t \to 1-$, and the limit is positive. Consequently,

$$\liminf_{t\to 1-} \frac{f(a)^2 - g_a(t)^2}{1-t} > 0,$$

and the statement follows. □

The lemma just proved means that 1 is an isolated point of $\mathfrak{G}_a$.

As an immediate corollary, we have the following theorem.

THEOREM 1.8.– *There exists $t_a^* < 1$ such that $\mathrm{P}\big(\xi_a \in (t_a^*, 1)\big) = 0$, and the distribution of ξ_a has an atom at 1, i.e. $\mathrm{P}(\xi_a = 1) > 0$. Consequently, $a(s) = z(1,s)$ for all $s \in [t_a^*, 1]$.*

Further, we prove that the distribution of ξ_a has no other atoms.

THEOREM 1.9.– *For any $t \in (0,1)$, $\mathrm{P}(\xi_a = t) = 0$. Consequently, $a \in C[0,1]$.*

PROOF.– We start by computing for $t \in (0,1)$

$$\begin{aligned}
a(t+) - a(t) &= \mathrm{E}[z(\xi_a, t) \mid \xi_a > t] - \mathrm{E}[z(\xi_a, t) \mid \xi_a \ge t] \\
&= \frac{\mathrm{E}[z(\xi_a,t)\mathbf{1}_{\xi_a > t}]\mathrm{P}(\xi_a \ge t) - \mathrm{E}[z(\xi_a,t)\mathbf{1}_{\xi_a \ge t}]\mathrm{P}(\xi_a > t)}{\mathrm{P}(\xi_a > t)\mathrm{P}(\xi_a \ge t)} \\
&= \frac{\mathrm{E}[z(\xi_a,t)\mathbf{1}_{\xi_a > t}]\mathrm{P}(\xi_a = t) - \mathrm{E}[z(\xi_a,t)\mathbf{1}_{\xi_a = t}]\mathrm{P}(\xi_a > t)}{\mathrm{P}(\xi_a > t)\mathrm{P}(\xi_a \ge t)} \\
&= \frac{\mathrm{E}[z(\xi_a,t)\mathbf{1}_{\xi_a > t}]\mathrm{P}(\xi_a = t)}{\mathrm{P}(\xi_a > t)\mathrm{P}(\xi_a \ge t)} - \frac{\mathrm{E}[z(t,t)\mathbf{1}_{\xi_a = t}]}{\mathrm{P}(\xi_a \ge t)} \\
&= \frac{a(t+)\mathrm{P}(\xi_a = t)}{\mathrm{P}(\xi_a \ge t)}.
\end{aligned} \tag{1.43}$$

Further, it is easy to see that $g = g_a$ has left and right derivatives at t, and

$$g_-'(t) = a(t)^2 + 2\int_0^t \big(z(t,s) - a(s)\big) z_t'(t,s)\, ds,$$

$$g_+'(t+) = a(t+)^2 + 2\int_0^t \big(z(t,s) - a(s)\big) z_t'(t,s)\, ds.$$

But for any $t \in \mathfrak{G}_a$, $g'_-(t) \geq 0$, $g'_+(t+) \leq 0$, so $a(t) \geq a(t+)$, whence from [1.43] we have that $a(t+) = a(t)$ and also $\mathrm{P}(\xi_a = t) = 0$, as $a(t+) > 0$. For $t \notin \mathfrak{G}_a$, $\mathrm{P}(\xi_a = t) = 0$ (recall that ξ_a takes values in $\mathfrak{G}_a$) and $a(t+) = a(t)$. $\square$

REMARK 1.6.– Due to monotonicity of z in the first variable, the right-hand side of inequality [1.19] is maximal for $t_2 = 1$, so we have that

$$f(a) \geq \frac{1}{4} \max_{t \in [0,1]} \int_0^t \big(z(1,s) - z(t,s)\big)^2 \, ds. \qquad [1.44]$$

Theorem 1.9 implies in particular that the inequality is strict, i.e. this lower bound is not reached. Indeed, if there were equality in [1.44], Lemma 1.5 would imply that the distribution of ξ_a is $\frac{1}{2}(\delta_{t_0} + \delta_1)$, where t_0 is the point where the minimum of the right-hand side of [1.44] is reached, which would contradict Theorem 1.9.

REMARK 1.7.– From [1.41] it is easy to see that the minimizing function $a = a(s)$ decreases on the intervals of the complement to $\mathfrak{G}_a$, particularly on $(0, t_1)$ and $(t_a^*, 1)$. The numerical experiments suggest that a decreases on the entire complement to $\mathfrak{G}_a$, except for $H = 0.9$, where the experiment gives $a(t_1) < a(t_a^*)$, but the difference $a(t_a^*) - a(t_1)$ is small and can be explained by the effect of discretization of the time.

1.8. Approximation of a discrete-time fBm by martingales

Obviously, in the previous sections we collected a lot of interesting properties of the minimizing function for the kernel z that corresponds to fBm. However, there is no clear analytical representation for this function. In this connection, let us give a numerical algorithm of its approximation. Of course, it is easier to consider the discrete-time approximations. So, in this section, we consider a problem of minimization of the principal functional, but in discrete time. This is an approximation to the original problem, so its solution can be considered as an approximate solution to the original problem. However, to be absolutely precise, we now simply solve a discrete analog of the minimization problem described above.

1.8.1. *Description of the discrete-time model*

We consider the discretizing procedure on the interval $[0, 1]$ for technical simplicity. Let N be a natural number, and define $b_k = B^H_{k/N}$, $k = 0, \ldots, N$. The vector $\boldsymbol{b} = (b_0, b_1, \ldots, b_N)$ will be called a discrete-time fBm. It generates a discrete-time filtration $\mathcal{F}_k = \sigma(b_0, \dots, b_k)$, $k = 0, \dots, N$. For an arbitrary random vector $\xi = (\xi_0, \xi_1, \dots, \xi_N)$ with square integrable components denote

$$G(\xi) = \max_{k=0,\dots,N} \mathrm{E}(b_k - \xi_k)^2.$$

We consider the problem of minimization of the functional $G(\xi)$, where ξ is an $\mathcal{F}_k$-martingale. Let us represent the vector $\boldsymbol{b}$ as a linear operator applied to a random vector with standard multivariate normal distribution. It follows that $b_0 = 0$ a.s. The vector $(b_1, \dots, b_N)^\top$ has multivariate normal distribution with zero mean and some covariance matrix Σ_b. The matrix Σ_b is non-singular (see Theorem 1.1), and its ijth element is equal to

$$(\Sigma_b)_{ij} = \frac{i^{2H} + j^{2H} - |i - j|^{2H}}{2N^{2H}}.$$

Since Σ_b is a positive definite, symmetric matrix, a decomposition exists:

$$\Sigma_b = KK^\top \qquad [1.45]$$

with lower-triangular K, and the matrix K is non-singular. The representation [1.45] is called the Cholesky decomposition. Cholesky decomposition is widely used for drawing a Gaussian random vector with specified covariance matrix. For the algorithm of computing the Cholesky decomposition and for its use in statistical simulations, see Section A3.1 in Appendix 3 and [CHA 15, Section 6.5.3]. We used `chol` function from R package. After applying Cholesky decomposition, we obtain the representation

$$(b_1, \dots, b_N)^\top = K(\zeta_1, \dots, \zeta_N)^\top,$$

where $\zeta_1, \dots, \zeta_N$ are independent random variables with standard normal distribution. Denote k_{ts} the elements of the matrix K, $1 \le t \le N$, $1 \le s \le N$. Since the matrix K is lower triangular, $k_{ts} = 0$ for $t < s$. Then,

$$b_t = \sum_{s=1}^{t} k_{ts} \zeta_s. \qquad [1.46]$$

Now compare the discrete-time representation [1.46] to the continuous-time Molchan representation [1.4]. We have $b_t = B^H_{t/N}$. Both ζ_s in [1.46] and $\sqrt{N}(W_{s/N} - W_{(s-1)/N})$ in [1.4] have a standard normal distribution. Thus, k_{ts} plays the similar role in [1.46] as the average value of the function $N^{-1/2} z\left(\frac{t}{N}, u\right)$ on interval $\left(\frac{s-1}{N}, \frac{s}{N}\right)$ in [1.4]. Therefore, the matrix K can be regarded as the discrete-time counterpart of the Molchan kernel $z(t, s)$.

Further, we will show, as in the continuous case, that minimization of G over martingales is equivalent to minimization over Gaussian martingales. Indeed, let $\xi = (\xi_0, \xi_1, \ldots, \xi_N)$, $\xi_0 = 0$, be an arbitrary square integrable $\mathcal{F}_k$-martingale. Owing to the fact that $\mathcal{F}_k = \sigma\{\zeta_1, \ldots, \zeta_k\}$, $k = 1, \ldots, N$, we have the following martingale representation:

$$\xi_n = \sum_{k=1}^{n} \alpha_k \zeta_k, \quad n = 1, \ldots, N,$$

where α_k is a square integrable $\mathcal{F}_{k-1}$-measurable random variable, $k = 1, \ldots, N$. Taking into account the martingale property of ξ and mutual independence of ζ_k, we obtain

$$\begin{aligned} G(\xi) &= \max_{j=0,\ldots,N} \mathrm{E}(b_j - \xi_j)^2 = \max_{j=1,\ldots,N} \sum_{n=1}^{j} \mathrm{E}(k_{jn} - \alpha_n)^2 \\ &= \max_{j=1,\ldots,N} \sum_{n=1}^{j} \left((k_{jn} - \mathrm{E}\alpha_n)^2 + \mathrm{Var}(\alpha_n) \right) \\ &\geq \max_{j=1,\ldots,N} \sum_{n=1}^{j} (k_{jn} - \mathrm{E}\alpha_n)^2. \end{aligned}$$

Hence, we can assume that ξ has a form $\xi_n = \sum_{k=1}^{n} a_k \zeta_k$, $n = 1, \ldots, N$, with some non-random $a_1, \ldots, a_n$. Then,

$$G(\xi) = \max_{t=1,\ldots,N} \sum_{s=1}^{t} (k_{ts} - a_s)^2 =: F(a). \quad [1.47]$$

Thus, we are searching for the minimum of functional $F(a)$ over $a \in \mathbb{R}^N$.

1.8.2. *Iterative minimization of the squared distance using alternating minimization method*

Now reduce the problem of finding the minimum of the convex functional $F(a)$ to the problem of finding the minimum of biconvex functional of two vectors. In order to introduce a functional of two vectors, denote

$$F(a, b) = \max_{t=1,\ldots,N} \left(\sum_{s=1}^{t} (a_s - k_{ts})^2 + \sum_{s=t+1}^{N} (a_s - b_s)^2 \right),$$

(Here, by convention, $\sum_{s=N+1}^{N}(a_s - b_s)^2 = 0$.)

For fixed a, $\min F(a,b)$ is reached at point $b = a$, i.e.

$$F(a) = F(a,a) = \min_{b\in\mathbb{R}^N} F(a,b). \qquad [1.48]$$

DEFINITION 1.2.– *Let Z be a non-empty bounded set in $\mathbb{R}^N$. Chebyshev center of the set Z is the point $a \in \mathbb{R}^N$ where the minimum of $\sup_{z\in Z}\|a - z\|$ is reached:*

$$a = \arg\min_{a\in\mathbb{R}^N}\sup_{z\in Z}\|a - z\| \qquad [1.49]$$

The minimum in [1.49] is reached because the criterion function

$$\sqrt{Q(a)} = \sup_{z\in Z}\|a - z\|$$

is continuous on $\mathbb{R}^N$ and tends to $+\infty$ as $\|a\| \to \infty$. It is reached at a unique point, according to the next lemma.

LEMMA 1.15.– *Square criterion function*

$$Q(a) = \sup_{z\in Z}\|a - z\|^2$$

is strongly convex.

PROOF.– To prove convexity, note that for every $z \in \mathbb{R}^N$, $a_1 \in \mathbb{R}^N$, $a_2 \in \mathbb{R}^N$, and $t \in (0,1)$

$$\begin{aligned}\|ta_1 + (1-t)a_2 - z\|^2 = t^2\|a_1\|^2 + 2t(1-t)\,a_1^\top a_2 + (1-t)^2\|a_2\|^2 \\ - 2ta_1^\top z - 2(1-t)a_2^\top z + \|z\|^2.\end{aligned}$$

This equality can be obtained easily by representing the square norm as the inner product of a vector by itself ($\|z\|^2 = z^\top z$) and using linearity and symmetry of the inner product. Indeed,

$$\begin{aligned}&\|ta_1 + (1-t)a_2 - z\|^2 = (ta_1 + (1-t)a_2 - z)^\top(ta_1 + (1-t)a_2 - z)\\ &= t^2a_1^\top a_1 + t(1-t)a_1^\top a_2 - ta_1^\top z\\ &\quad + (1-t)ta_2^\top a_1 + (1-t)^2a_2^\top a_2 - (1-t)a_2^\top z\\ &\quad - tz^\top a_1 - (1-t)z^\top a_2 + z^\top z\end{aligned}$$

$$
\begin{aligned}
&= (t - t(1-t))a_1^\top a_1 + t(1-t)a_1^\top a_2 - t a_1^\top z \\
&\quad + t(1-t)a_2^\top a_1 + ((1-t) - t(1-t))a_2^\top a_2 - (1-t)a_2^\top z \\
&\quad - t z^\top a_1 - (1-t)z^\top a_2 + (t + (1-t))z^\top z \\
&= t a_1^\top a_1 - t(1-t)a_1^\top a_1 + t(1-t)a_1^\top a_2 - t a_1^\top z \\
&\quad + t(1-t)a_2^\top a_1 + (1-t)a_2^\top a_2 - t(1-t)a_2^\top a_2 - (1-t)a_2^\top z \\
&\quad - t z^\top a_1 - (1-t)z^\top a_2 + t z^\top z + (1-t)z^\top z \\
&= t a_1^\top a_1 - t a_1^\top z - t z^\top a_1 + t z^\top z \\
&\quad + (1-t)a_2^\top a_2 - (1-t)a_2^\top z - (1-t)z^\top a_2 + (1-t)z^\top z \\
&\quad - t(1-t)a_1^\top a_1 + t(1-t)a_1^\top a_2 + t(1-t)a_2^\top a_1 - t(1-t)a_2^\top a_2 \\
&= t(a_1 - z)^\top (a_1 - z) \\
&\quad + (1-t)(a_2 - z)^\top (a_2 - z) \\
&\quad - t(1-t)(a_1 - a_2)^\top (a_1 - a_2) \\
&= t\,\|a_1 - z\|^2 + (1-t)\,\|a_2 - z\|^2 - t(1-t)\,\|a_1 - a_2\|^2. \qquad \square
\end{aligned}
$$

If, in addition to assumptions $a_1 \in \mathbb{R}^N$, $a_2 \in \mathbb{R}^N$ and $t \in (0,1)$, the inequality $a_1 \neq a_2$ holds true, then

$$
\begin{aligned}
Q(ta_1 + (1-t)\,a_2) &\le tQ(a_1) + (1-t)Q(a_2) - t(1-t)\|a_1 - a_2\|^2 \\
&< tQ(a_1) + (1-t)Q(a_2).
\end{aligned}
$$

Thus, the square criterion function $Q(a)$ is indeed strongly convex.

COROLLARY 1.6.– *As a consequence, square criterion function $Q(a)$ cannot reach its minimum at more than one point.*

REMARK 1.8.– Lemma 1.15 can be generalized for any Hilbert space. Namely, for $\mathcal{H}$ the Hilbert space, for any fixed element $z \in \mathcal{H}$, the functional $Q(a) = \|a - z\|^2$ is strictly convex. The proof is exactly the same if only the notation $a^\top b$ is used instead of the standard notation of the scalar product in $\mathcal{H}$.

For fixed b, $\min F(a,b)$ is reached at point a being the Chebyshev center of the N-point set $\{k_1(b), k_2(b), \ldots, k_N(b)\}$, where

$$k_t(b) = \begin{pmatrix} k_{t1} \\ \vdots \\ k_{tt} \\ b_{t+1} \\ \vdots \\ b_N \end{pmatrix}, \quad t = 1, \ldots, N-1; \quad k_N(b) = \begin{pmatrix} k_{N1} \\ k_{N2} \\ k_{N3} \\ \vdots \\ k_{N,N-1} \\ k_{NN} \end{pmatrix} = k_{N\bullet}. \quad [1.50]$$

Due to [1.48], the problem of minimization of $F(a)$ is equivalent to the problem of minimization of $F(a,b)$:

$$\min_{a \in \mathbb{R}^N} F(a) = \min_{a \in \mathbb{R}^N} \min_{b \in \mathbb{R}^N} F(a,b).$$

If the minimum of $F(a)$ is reached at point $a = a_*$, then the minimum of $F(a,b)$ is reached at point $a = b = a_*$. If the minimum of $F(a,b)$ is reached at point $a = a_*$, $b = b_*$, then the minimum of $F(a)$ is reached at point $a = a_*$. In what follows, $\min F$ will mean a common minimum of functions $F(a)$ and $F(a,b)$.

Summarize the properties of the functions $F(a)$ and $F(a,b)$ in the following proposition.

PROPOSITION 1.1.–

1) For a fixed $a \in \mathbb{R}^N$, the minimum $\min_{b \in \mathbb{R}^N} F(a,b) = F(a)$ is reached at point $b = a$.

2) For a fixed $b \in \mathbb{R}^N$, the minimum $\min_{a \in \mathbb{R}^N} F(a,b)$ is reached at point a being the Chebyshev center of the points $k_1(b), k_2(b), \ldots, k_N(b)$.

3) The minimal values of the functions $F(a,b)$ and $F(a)$ coincide. We denote their common value by $\min F$:

$$\min F := \min_{a,b \in \mathbb{R}^N} F(a,b) = \min_{a \in \mathbb{R}^N} F(a).$$

Now, we are in position to find the minimum of $F(a,b)$ by the procedure that is called an alternating minimization. Namely, let $a^{(0)} \in \mathbb{R}^N$ be the initial approximation. In our case we take $k_{N\bullet}$ as $a^{(0)}$. Then, the minimum of $F(a^{(0)}, b)$ is reached at point $b = a^{(0)}$, and we minimize $F(a, a^{(0)})$ with respect to a:

$$a^{(1)} = \arg\min_{a \in \mathbb{R}^N} F(a, a^{(0)}).$$

Again, the minimum of $F(a^{(1)}, b)$ is reached at point $b = a^{(1)}$. Define the sequence $\{a^{(n)}, n \geq 1\}$ iteratively:

$$a^{(n)} = \arg\min_{a \in \mathbb{R}^N} F(a, a^{(n-1)}), \quad n \geq 1. \qquad [1.51]$$

Then, we have that the criterion function is non-increasing:

$$\begin{aligned} F(a^{(0)}, a^{(0)}) &\geq F(a^{(1)}, a^{(0)}) \geq F(a^{(1)}, a^{(1)}) \\ &\geq F(a^{(2)}, a^{(1)}) \geq F(a^{(2)}, a^{(2)}) \geq \ldots, \end{aligned} \qquad [1.52]$$

and since $F(a^{(k)}, a^{(k)}) = F(a^{(k)})$,

$$F(a^{(0)}) \geq F(a^{(1)}) \geq F(a^{(2)}) \geq \ldots, \qquad [1.53]$$

According to the following theorem, the sequence $\{F(a^{(k)}), \ k \geq 0\}$ converges to $\min F$.

THEOREM 1.10.– *Let*

$$\min_{t=s,\ldots,N} k_{ts} \leq a_s^{(0)} \leq \max_{t=s,\ldots,N} k_{ts} \quad \textit{for all } s,\ 1 \leq s \leq N. \qquad [1.54]$$

Then, the sequence $\{a^{(n)}, \ n \geq 1\}$ defined by [1.51] *has the following properties:*

1) $\lim_{n\to\infty} F(a^{(n)}) = \min F$.

2) If the minimal value of $F(a)$ is reached at the unique point a_* (i.e. the equality $F(a) = \min F$ holds true if and only if $a = a_*$), then

$$\lim_{n\to\infty} a^{(n)} = a_*.$$

PROOF.– 1) Introduce a following rectangle $H \subset \mathbb{R}^N$:

$$\begin{aligned} H &= \{a \in \mathbb{R}^N \mid \forall s = 1, \ldots, N : \min_{t=s,\ldots,N} k_{ts} \leq a_s \leq \max_{t=s,\ldots,N} k_{ts}\} \\ &= \left[\min_{t=1,\ldots,N} k_{t1}, \max_{t=1,\ldots,N} k_{t1}\right] \times \left[\min_{t=2,\ldots,N} k_{t2}, \max_{t=2,\ldots,N} k_{t2}\right] \\ &\quad \times \ldots \times [\min(k_{N-1,N-1}, k_{N,N-1}), \max(k_{N-1,N-1}, k_{N,N-1})] \\ &\quad \times \{k_{NN}\}. \end{aligned} \qquad [1.55]$$

Denote by d the length of the diagonal of the rectangle H, so that

$$d^2 = \sum_{j=1}^{N} \left(\max_{i=j,\ldots,N} k_{ij} - \min_{i=j,\ldots,N} k_{ij} \right)^2 . \qquad [1.56]$$

Condition [1.54] means that $a^{(0)} \in H$. By induction, $a^{(n)} \in H$ for all n. Indeed, if $a^{(n-1)} \in H$, then $k_t(a^{(n-1)}) \in H$ for all $t = 1, \ldots, N$, where $k_t(b)$ is defined in [1.50]. The Chebyshev center of a set lies in the convex hull of the set. Thus, $a^{(n)}$ lies in the convex hull of $k_i(a^{(n-1)})$, $i = 1, \ldots, N$, whence $a^{(n)} \in H$.

The minimal value $\min_{a\in\mathbb{H}} F(a)$ is reached because $F(a)$ is a continuous function and H is a non-empty closed bounded set. Denote by $a_* \in H$ the point where the minimal value $\min_{a\in\mathbb{H}} F(a)$ is reached. It holds that for all $a \in \mathbb{R}^N$

$$F(a) \geq F(\text{nearest}(H, a)) \geq F(a_*), \qquad [1.57]$$

whence $F(a_*) = \min_{a\in\mathbb{R}^N} F(a)$. Here, nearest$(H, a)$ is the point of H nearest to a; the coordinates of nearest(H, a) equal

$$(\text{nearest}(H, a))_s = \begin{cases} \min\limits_{t=s,\ldots,N} k_{ts}, & \text{if} \quad a_s \leq \min_{t=s,\ldots,N} k_{ts}, \\ a_s, & \text{if} \quad \min_{t=s,\ldots,N} k_{ts} \leq a_s \\ & \leq \max_{t=s,\ldots,N} k_{ts}, \\ \max\limits_{t=s,\ldots,N} k_{ts}, & \text{if} \quad a_s \geq \max_{i=j,\ldots,N} k_{ts}. \end{cases}$$

Due to inequality [1.57], the minimal value of $F(a)$ on $\mathbb{R}^N$ is reached at a_*, namely

$$\min F = F(a_*) = \min_{a\in\mathbb{R}^N} F(a) = \min_{a,b\in\mathbb{R}^N} F(a, b). \qquad [1.58]$$

Since $a^{(n)} \in H$ and $a_* \in H$, we have $\|a^{(n)} - a_*\| \leq d$ for all $n \geq 0$. For all $a \in H$ and $b \in H$, the inequality $F(a) \leq F(a, b) \leq d^2$ holds true.

The case $d = 0$ is trivial: in this case, the set H consists of only one point, $a^{(n)} = a^{(0)}$ and $F(a^{(n)}) = 0$ for all n; therefore, the statement of this theorem holds true. Hence, in the rest of the proof assume that $d > 0$.

Define the following pseudonorms on $\mathbb{R}^N$:

$$\|x\|_t = \left(\sum_{s=1}^{t} x_s^2 \right)^{1/2} \qquad \text{and} \qquad \|x\|_{\perp t} = \left(\sum_{s=t+1}^{N} x_s^2 \right)^{1/2} .$$

For $t = N$ we have that $\|x\|_N = \|x\|$ and $\|x\|_{\perp N} = 0$. With this notation,

$$\|x\|_t^2 + \|x\|_{\perp t}^2 = \|x\|^2 \quad \text{for all } t = 1, \dots, N \text{ and } x \in \mathbb{R}^N,$$

$$F(a) = \max_{t=1,\dots,N} \|a - k_{t\bullet}\|_t^2, \tag{1.59}$$

$$F(a,b) = \max_{t=1,\dots,N} \left(\|a - k_{i\bullet}\|_t^2 + \|a - b\|_{\perp t}^2\right).$$

In what follows, we are going to use inequalities

$$\|a - k_{i\bullet}\|_i^2 \le F(a) \quad \text{for all } a \in \mathbb{R}^N,$$

$$\|a - b\|_{\perp i}^2 \le F(a,b) \le d^2 \quad \text{for all } a \in H \text{ and } b \in H.$$

Now we are going to construct an upper bound for $F(a^{(n)})$, and it will be inequality [1.62].

Denote for fixed n

$$\alpha_n = \frac{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)\sqrt{F(a^{(n-1)})}}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}.$$

Then,

$$1 - \alpha_n = \frac{d^2 - \left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)\sqrt{F(a_*)}}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}.$$

Taking into account the relations $0 \le F(a_*) \le F(a^{(n-1)}) \le d^2$, we obtain the inequality $0 \le \alpha_n < 1$.

The next auxiliary result also will be applied to obtain [1.62]. Namely, we construct the upper bound for $F((1-\alpha_n)a^{(n-1)} + \alpha_n a_*,\ a^{(n-1)})$. In order to do this, note that for every $t = 1, \dots, N$,

$$\begin{aligned}
&\|(1-\alpha_n)a^{(n-1)} + \alpha_n a_* - k_{i\bullet}\|_t \\
&\quad \le (1-\alpha_n)\|a^{(n-1)} - k_{i\bullet}\|_i + \alpha_n \|a_* - k_{i\bullet}\|_t \\
&\quad \le (1-\alpha_n)\sqrt{F(a^{(n-1)})} + \alpha_n \sqrt{F(a_*)}
\end{aligned}$$

$$= \frac{d^2 \sqrt{F(a^{(n-1)})}}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2},$$

and

$$\|(1-\alpha_n)a^{(n-1)} + \alpha_n a_* - a^{(n-1)}\|_{\perp t} = \alpha_n \|a_* - a^{(n-1)}\|_{\perp t} \le \alpha_n d$$
$$= \frac{d\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right) \sqrt{F(a^{(n-1)})}}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}.$$

Hence,

$$\|(1-\alpha_n)a^{(n-1)} + \alpha_n a_* - k_{i\bullet}\|_t^2 + \|(1-\alpha_n)a^{(n-1)} + \alpha_n a_* - a^{(n-1)}\|_{\perp t}^2$$
$$\le \frac{d^4 F(a^{(n-1)}) + d^2 \left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 F(a^{(n-1)})}{\left(\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2\right)^2}$$
$$= \frac{d^2 F(a^{(n-1)})}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}. \qquad [1.60]$$

Take the maximum over $t = 1, \dots, N$ in the left-hand side of [1.60], apply equation [1.59], and obtain

$$F((1-\alpha_n)a^{(n-1)} + \alpha_n a_*,\ a^{(n-1)}) \le \frac{d^2 F(a^{(n-1)})}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}.$$

Now we continue with the upper bound for $F(a^{(n)})$. Recall that $F(a) = \min_b F(a, b)$ and $F(a^{(n)}, a^{(n-1)}) = \min_a F(a, a^{(n-1)})$. Therefore, the following inequality holds true:

$$F(a^{(n)}) \le F(a^{(n)}, a^{(n-1)}) \le F((1-\alpha_n)a^{(n-1)} + \alpha_n a_*,\ a^{(n-1)})$$
$$\le \frac{d^2 F(a^{(n-1)})}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2} \le F(a^{(n-1)}). \qquad [1.61]$$

The sequence $\{F(a^{(n)}),\ n \geq 0\}$ is decreasing and bounded, more exactly, $0 \leq F(a^{(n)}) \leq F(a^{(0)})$. Hence, it converges to a finite limit. From inequality [1.61], we finally obtain the desired upper bound:

$$F(a^{(n)}) \leq \frac{d^2 F(a^{(n-1)})}{\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 + d^2}. \qquad [1.62]$$

Take the limit in [1.62] as $n \to \infty$:

$$\lim_{n\to\infty} F(a^{(n)}) \leq \frac{d^2 \lim_{n\to\infty} F(a^{(n)})}{\left(\sqrt{\lim_{n\to\infty} F(a^{(n)})} - \sqrt{F(a_*)}\right)^2 + d^2}.$$

Therefore,

$$\left(\sqrt{\lim_{n\to\infty} F(a^{(n)})} - \sqrt{F(a_*)}\right)^2 \lim_{n\to\infty} F(a^{(n)}) \leq 0,$$

whence either $\sqrt{\lim_{n\to\infty} F(a^{(n)})} - \sqrt{F(a_*)} = 0$, or $\lim_{n\to\infty} F(a^{(n)}) \leq 0$.

If $\sqrt{\lim_{n\to\infty} F(a^{(n)})} - \sqrt{F(a_*)} = 0$, then $\lim_{n\to\infty} F(a^{(n)}) = F(a_*)$. Since $F(a_*) = \min_a F(a) \geq 0$, the inequality $0 \leq F(a_*) \leq \lim_{n\to\infty} F(a^{(n)})$ holds true, and if $\lim_{n\to\infty} F(a^{(n)}) \leq 0$, then $F(a_*) = \lim_{n\to\infty} F(a^{(n)}) = 0$. Thus, the equality $\lim_{n\to\infty} F(a^{(n)}) = F(a_*)$ holds true in either case. The first statement of the theorem is proved.

2) Let us draw the reader's attention to the fact that for the second statement of the theorem, we have an additional condition, namely it is assumed that the function $F(a)$ reaches its minimum at the unique point a_*. In the proof of the first part, we established that $F(a)$ reaches its minimum in H. Thus, $a_* \in H$. Fix $\epsilon > 0$ and prove that $\|a^{(n)} - a_*\| < \epsilon$ for n large enough. Denote

$$H_\epsilon = \{a \in H : \|a^{(n)} - a_*\| \geq \epsilon\}.$$

The set $H_\epsilon \subset H \subset \mathbb{R}^N$ is closed and bounded. Recall that $a^{(n)} \in H$ for all n. If $H_\epsilon = \emptyset$, then $\|a^{(n)} - a_*\| < \epsilon$ for all n. If $H_\epsilon \neq \emptyset$, then the function $F(a)$, as a continuous function on a non-empty compact set, reaches its minimum on H_ϵ, and

$$\min_{a\in H_\epsilon} F(a) > \min F = F(a_*) = \lim_{n\to\infty} F(a^{(n)}),$$

since $a_* \notin H_\epsilon$ and $F(a)$ reaches its minimal value $\min F$ at the unique point, which necessarily must coincide with a_*.

Therefore,

$$\min_{a \in H_\epsilon} F(a) > F(a^{(n)})$$

for n large enough, whence

$$a^{(n)} \notin H_\epsilon, \qquad a^{(n)} \in H \setminus H_\epsilon, \qquad \|a^{(n)} - a_*\| < \epsilon$$

for n large enough.

We proved that for any $\epsilon > 0$, n_0 exists such that for $n \geq n_0$ the inequality $\|a^{(n)} - a_*\| < \epsilon$ holds true. Then $a^{(n)} \to a_*$. The second statement of this theorem is proved. □

REMARK 1.9.– Generally speaking, condition [1.54] in Theorem 1.10 can be omitted. In order to prove Theorem 1.10 without condition [1.54], we can dismiss definitions [1.55] and [1.56] and define H and d as follows. Let H be the axis-aligned minimum bounding box that contains the points $a^{(0)}$, $k_1(a^{(0)})$, $k_2(a^{(0)}), \ldots,$ and $k_N(a^{(0)})$. In other words, we extend H in such a way that $a^{(0)} \in H$. Then, let d be the length of the diagonal of H. With these definitions, and without condition [1.54], the further proof remains correct.

REMARK 1.10.– 1) Theorem 1.10 is valid for arbitrary matrix $K = (k_{ts})_{t,s=1}^N$. The fact that K is a factor of the Cholesky decomposition of the covariance matrix Σ_b, is not used in the proof.

2) In Theorem 1.5, we proved that continuous-time functional $f(a)$, $a \in L_2([0,1])$, has a unique minimum. We have not proved a similar result for discrete-time functional $F(a)$, $a \in \mathbb{R}^N$.

REMARK 1.11.– From inequality [1.62], we obtain the following bound:

$$\left(\sqrt{F(a^{(n-1)})} - \sqrt{F(a_*)}\right)^2 \leq \frac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}.$$

Hence,

$$0 \le F(a^{(n-1)}) - F(a_*)$$
$$\le \begin{cases} 2d\sqrt{\dfrac{F(a^{(n-1)})(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}} - \dfrac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}, \\ \qquad \text{if} \quad F(a^{(n-1)}) \ge \dfrac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}, \\ \dfrac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}, \\ \qquad \text{if} \quad F(a^{(n-1)}) \le \dfrac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}. \end{cases} \qquad [1.63]$$

Finally, it follows from [1.63] that

$$0 \le F(a^{(n-1)}) - F(a_*)$$
$$\le 2d\sqrt{\frac{F(a_*)(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}} + \frac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}$$
$$\le 2d\sqrt{F(a^{(n-1)}) - F(a^{(n)})} + \frac{d^2(F(a^{(n-1)}) - F(a^{(n)}))}{F(a^{(n)})}.$$

Using inequalities $F(a^{(0)}) \le d^2$ and [1.62], we can construct the following bounds. If $F(a_*) = 0$, then $0 \le F(a^{(n)}) \le d^2/(n+1)$. Otherwise, if $F(a_*) > 0$, then

$$F(a_*) \le F(a^{(n)}) \le F(a_*) + \frac{d^4}{4nF(a_*)}.$$

1.8.3. *Implementation of the alternating minimization algorithm*

We approximate the vector that minimizes the functional $F(a)$ as follows. As the initial approximation, we take the bottom row of the matrix K, i.e. $a^{(0)} = k_{N\bullet}$. Then we iteratively perform minimization in [1.51] using the Chebyshev center algorithm, which is presented in Appendix 2. We stop iterations [1.51] when both conditions become true: the upper bound [1.63] is lower than the threshold and the distance $\|a^{(n-1)} - a^{(n)}\|$ is less than the threshold. Finally, we take $a^{(n)}$ as an approximation of the point of minimum.

The convergence of the algorithm is demonstrated in the following experiment. We took $H = 0.6$ and $N = 1000$. We performed 10 iterations of the algorithm and obtained 10 approximations $a^{(n)}$, $n = 1, \ldots, 10$ of the point a_* in which the function $F(a)$ reaches minimum, taking $a^{(0)} = k_{N\bullet}$ as the initial approximation. In Table 1.1 we present the following items: values of criterion function at $a^{(n)}$, $n = 1, \ldots, 10$, the right-hand side of inequality [1.63], which serves as the guaranteed upper bound for $F(a^{(n)}) - \min F$, and distances between some $a^{(n)}$'s.

n	$F(a^{(n)})$	Upper bound for $F(a^{(n)}) - \min F$	$\lVert a^{(n)} - a^{(n-1)} \rVert$	$\lVert a^{(n)} - a^{(10)} \rVert$
0	0.015646356	378.0	—	7.154×10^{-2}
1	0.006397266	62.89	7.495×10^{-2}	1.943×10^{-2}
2	0.005156821	2.006	1.915×10^{-2}	3.958×10^{-4}
3	0.005117567	2.426×10^{-3}	3.961×10^{-4}	1.286×10^{-6}
4	0.005117560	1.956×10^{-6}	1.284×10^{-6}	1.479×10^{-9}
5	0.005117560	—	1.477×10^{-9}	2.022×10^{-12}
6	0.005117560	7.377×10^{-8}	2.019×10^{-12}	6.353×10^{-14}
7	0.005117560	—	5.969×10^{-14}	5.938×10^{-14}
8	0.005117560	1.594×10^{-7}	6.227×10^{-14}	6.527×10^{-14}
9	0.005117560	1.413×10^{-7}	6.261×10^{-14}	6.735×10^{-14}
10	0.005117560	—	6.735×10^{-14}	0

Table 1.1. *Iterative minimization of the function $F(a)$: the value of criterion function $F(a^{(n)})$, the upper bound for $F(a^{(n)}) - \min F$ obtained from inequality [1.63], the distances between consecutive approximation to the point of minimum and between each approximation and the last approximation*

We can observe that the inequality $F(a^{(6)}) < F(a^{(5)})$, which should hold true theoretically, actually does not hold true due to rounding errors. The upper bound for $F(a^{(n)}) - \min F$ is very loose for $n \leq 4$, and does not make sense for those values n that do not meet the inequality $F(a^{(n+1)}) < F(a^{(n)})$. The approximations $a^{(n)}$ of the minimum point a_* seem to converge up to machine precision.

1.8.4. *Computation of the minimizing function*

Table 1.2 gives the minimum values of the functional $F(a)$ for various H and N. We tried $N = 100$, $N = 200$ and $N = 1000$ to see if the difference of N can explain the discrepancy between our results and those of [SHK 14]. The results in [SHK 14] for $N = 200$ are close to our results for $N = 100$.

N	H	.51	.55	.6	.65	.7	.75	.8	.85	.9	.95	.99
100	$\min F(a)$	.00005	.0013	.0051	.0112	.0200	.0320	.0482	.0704	.1022	.1511	.2190
200	$\min F(a)$	.00005	.0013	.0051	.0113	.0201	.0321	.0483	.0706	.1025	.1515	.2193
1000	$\min F(a)$	.00006	.0013	.0051	.0113	.0202	.0322	.0485	.0708	.1027	.1518	.2196

Table 1.2. *Values of* $\min F$ *for various* H *from* 0.51 *to* 0.99 *and for* $N = 100$*,* $N = 200$*, and* $N = 1000$

Figure 1.2 shows the values of $\min F$ for H from 0.51 to 0.99 with a step 0.010, and also for $H = 0.501, 0.995$ and 0.999, for $N = 1000$.

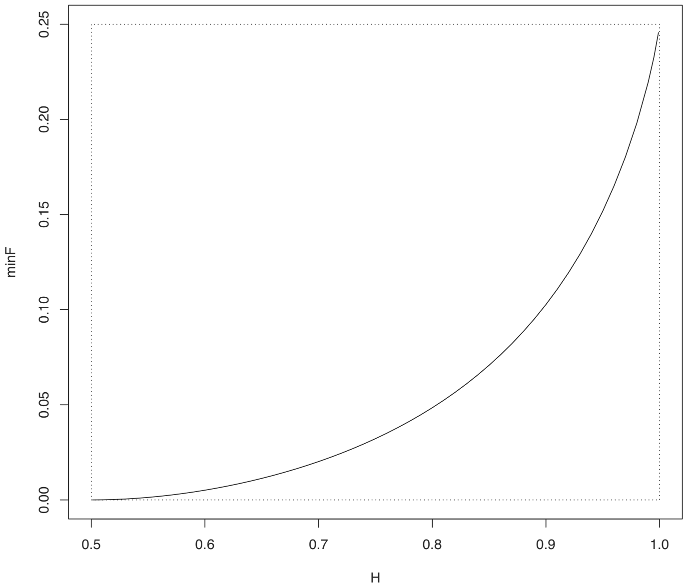


Figure 1.2. *Values of* $\min F$ *for* H *from* 0.501 *to* 0.999 *with step* 0.010 *and for* $H = 0.501, 0.995$ *and* 0.999

As a by-product, the Chebyshev center algorithm computes weights w_t in the linear combination

$$a = \sum_{i=1}^{N} w_t k_{t\bullet},$$

which can be used to obtain the distribution of the random variable ξ_a introduced in Lemma 1.12. Our numerical experiment shows that the random variable ξ_a is concentrated in the interval (t_1, t_a^*) and at point 1; thus

$\mathfrak{G}_a = [t_1, t_a^*] \cup \{1\}$. The support of ξ_a and $\mathrm{P}[\xi_a = 1]$ for $H = 0.51, 0.52, \ldots, 0.99$ are shown on Figure 1.3. For instance, for $H = 0.6$ the random variable ξ_a is concentrated in the interval $(0.8, 0.868)$ and at point 1, with

$$\mathrm{P}[\xi_a \in (0.8, 0.868)] = 0.693, \qquad [1.64]$$

and $\mathrm{P}[\xi_a = 1] = 0.307$. For $H = 0.75$, the random variable ξ_a is concentrated in $(0.080, 0.868)$ and at 1, with

$$\mathrm{P}[\xi_a \in (0.080, 0.868)] = 0.583, \qquad [1.65]$$

and $\mathrm{P}[\xi_a = 1] = 0.417$. For $H = 0.9$, ξ_a is concentrated in $(0.035, 0.208)$ and at 1, with

$$\mathrm{P}[\xi_a \in (0.035, 0.208)] = 0.52, \qquad [1.66]$$

and $\mathrm{P}[\xi_a = 1] = 0.48$. The distribution function of the continuous part of the distribution of ξ_a is shown on Figures 1.6, 1.7 and 1.8 for $H = 0.6$, $H = 0.75$ and $H = 0.9$, respectively.

We numerically evaluate the minimizing function for some $H < 0.5$, namely for $H = 0.1$ and $H = 0.4$, see Figures 1.4 and 1.5. The plots suggest that for $0 < H < 1/2$ the minimizing function $a(s)$ is equal to $\mathrm{E}[z(\xi_a, s) \mid \xi_a \geq s]$, (i.e. Theorem 1.7 holds true), where the random variable ξ_a satisfies the following properties:

1) the support of ξ_a is an interval $[t_1, 1]$ for some $t_1 \in (0, 1)$;

2) the distribution of ξ_a is absolutely continuous. Particularly, $\mathrm{P}(\xi_a = 1) = 0$;

3) the probability density function (pdf) $p_{\xi_a} = p_{\xi_a}(t), t \in [0, 1]$ of ξ_a satisfies the relation $\lim_{t \to 1-} p_{\xi_a}(t) = +\infty$.

However, this is an observation rather than a theoretically proven statement.

Figures 1.4–1.8 contain graphs of the minimizing vector (thick solid line) and the scaled squared distance $R(t) = \sum_{s=1}^{t} (k_{ts} - a_s)^2$ (thin solid line) for $H = 0.1, 0.4, 0.6, 0.75, 0.9$ and $N = 1000$. The key point of the whole theory we have constructed is that we consider the minimizing function $a = a(t)$ as a result of successive constructions of the vector that minimizes the functional $F = F(a)$, defined by equality [1.47]. As a result of the plotting, the minimizing function $a(t)$ varies slightly on the interval where $R(t)$ reaches its maximum, and decreases outside the interval.

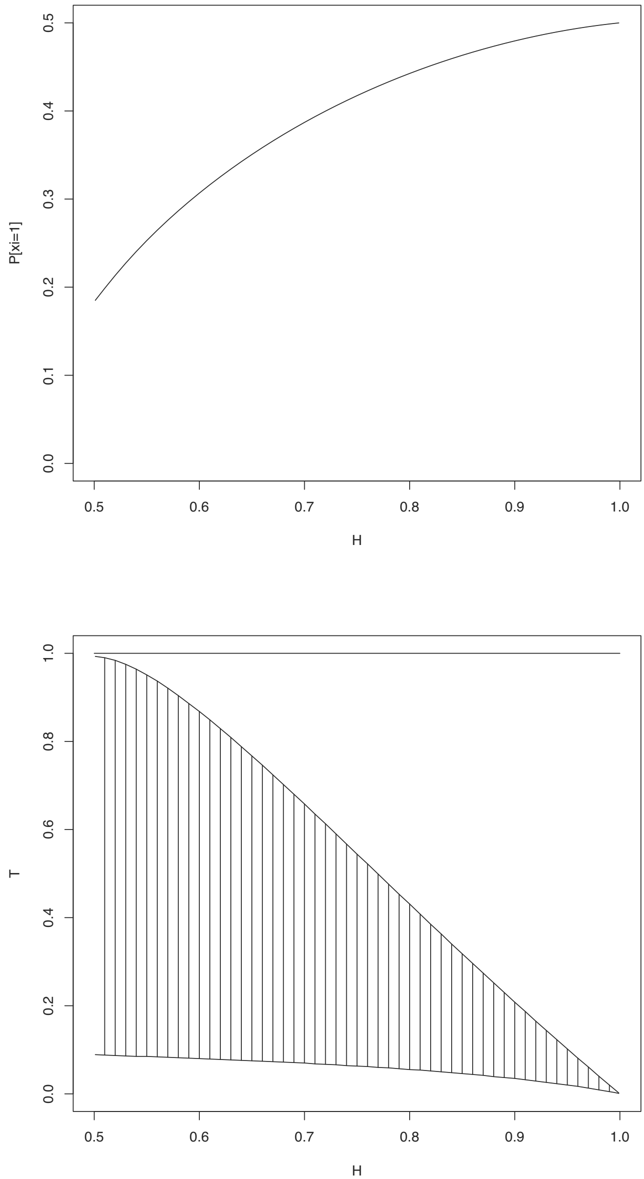


Figure 1.3. *Graph of* $\mathrm{P}[\xi_a = 1]$ *and the support of the random variable* ξ_a *for* H *from* 0.51 *to* 0.99 *with step* 0.010*.* (H, T) *is the point from the shaded area that is the support of* ξ_a

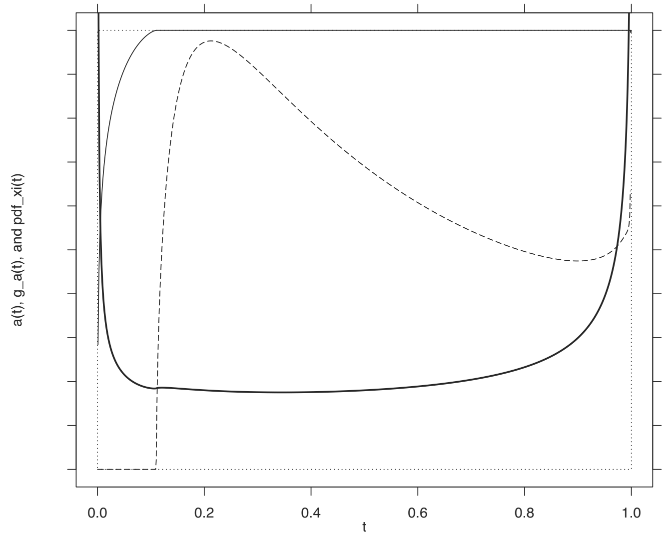


Figure 1.4. *The minimizing function* $a(t)$ *(thick solid line), the scaled square distance* $g_a(t)$ *(thin solid line), and the pdf of* ξ_a *(dashed line) for* $H = 0.1$

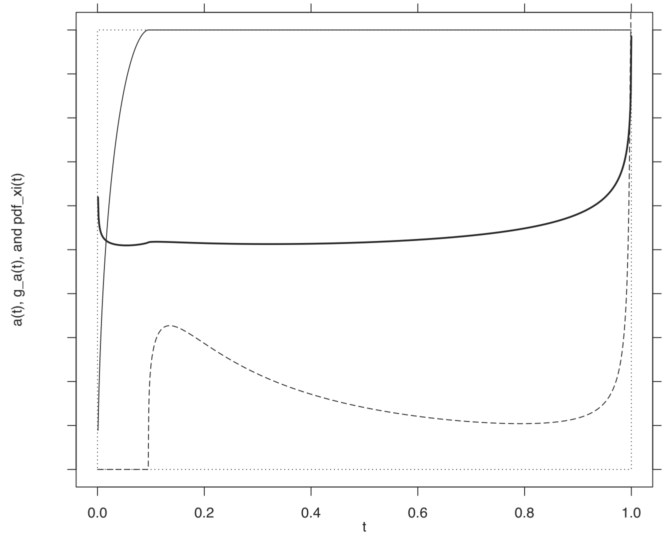


Figure 1.5. *The minimizing function* $a(t)$*, the scaled square distance* $g_a(t)$*, and the pdf of* ξ_a *for* $H = 0.4$

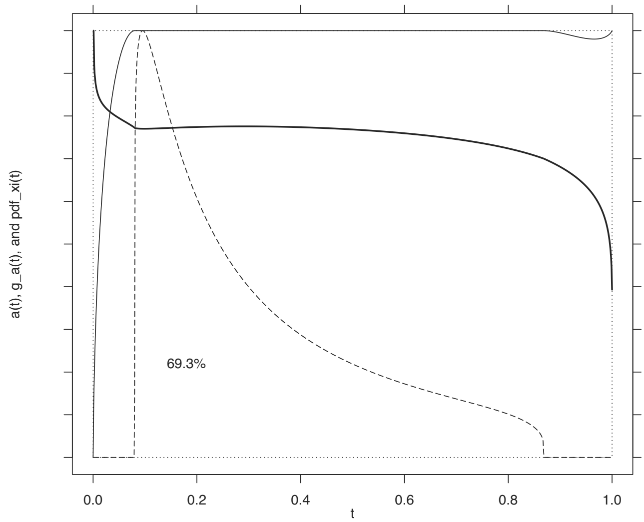


Figure 1.6. *The minimizing function $a(t)$, the scaled square distance $g_a(t)$, and the pdf of ξ_a for $H = 0.6$. Percentage corresponds to equality* $[1.64]$

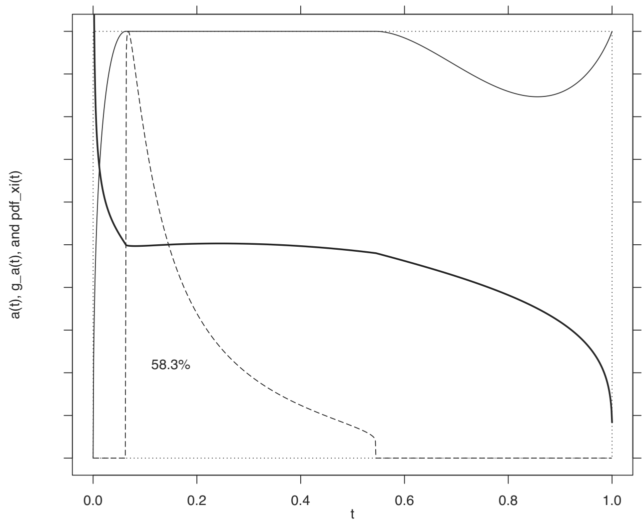


Figure 1.7. *The minimizing function $a(t)$, the scaled square distance, and the pdf of ξ_a for $H = 0.75$. Percentage corresponds to equality* $[1.65]$

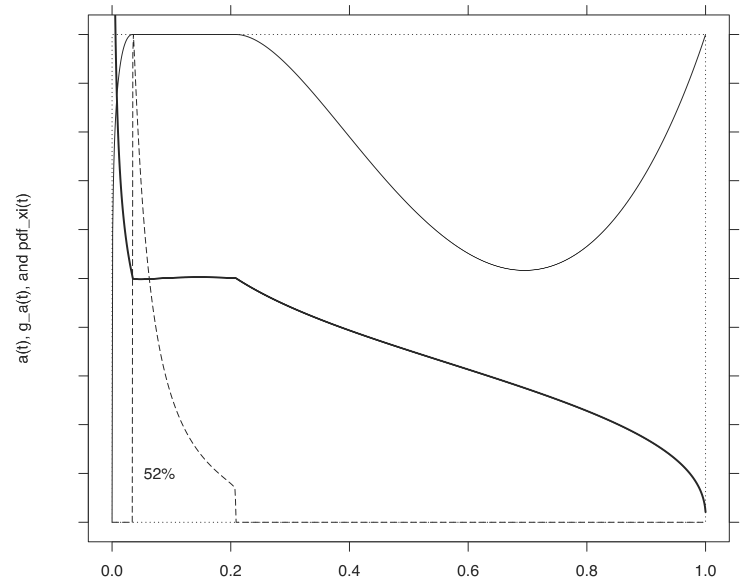


Figure 1.8. *The minimizing function $a(t)$, the scaled distance $g_a(t)$, and the pdf of ξ_a for $H = 0.9$. Percentage corresponds to equality* [1.66]

1.9. Exercises

EXERCISE 1.1.– Let $z : [0,1]^2 \to \mathbb{R}$ be a function such that

$$\sup_{t\in[0,1]} \int_0^t z(t,s)^2\, ds < \infty.$$

(Here we do not assume that z is the Molchan kernel of fBm.) For $a \in L_2([0,1])$ define functions

$$g_a(t) = \left(\int_0^t (a(s) - z(t,s))^2\, ds\right)^2, \qquad t \in [0,1]; \tag{1.67}$$

$$f(a) = \sup_{t\in[0,1]} g_a(t). \tag{1.68}$$

Prove the following statements:

1) The functional f reaches its minimal value on $L_2([0,1])$.

2) Suppose that for every function $a \in L_2([0,1])$ where $f(a)$ reaches minimum, $g_a(1) = f(a)$. Then the functional f reaches minimum at "essentially" unique point, i.e. if $f(a_1) = f(a_2) = \min f$, then the functions a_1 and a_2 are equal a.e. on $[0,1]$.

EXERCISE 1.2.– Consider the Riemann–Liouville process

$$R^\beta(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1}\, dW_s$$

for $\beta > \frac{1}{2}$. In this representation, which is similar to [1.4], the kernel is equal to

$$z(t,s) = \begin{cases} \frac{1}{\Gamma(\beta)}(t-s)^{\beta-1}, & \text{if } t < s, \\ 0, & \text{if } t \geq s. \end{cases}$$

Prove the following statements:

1) The Riemann–Liouville process, considered for $\beta > \frac{1}{2}$, is continuous in the mean-square sense, i.e.

$$\lim_{t_2 \to t_1} \mathrm{E}(R^\beta(t_2) - R^\beta(t_1))^2 = 0.$$

2) The functional $f(a)$ that is constructed in [1.67]–[1.68] for this $z(t,s)$, reaches its minimal value at "essentially" unique point.

EXERCISE 1.3.– Give an example of such a lower-triangular $N \times N$ matrix K (with entries k_{ts}, $1 \leq t \leq N$, $1 \leq s \leq N$) that the functional

$$F(a) = \max_{t=1,\ldots,N} \sum_{s=1}^{t} (a_s - k_{ts})^2$$

reaches its minimum at multiple points.