



HAL
open science

Probability density function of the local score position

Agnes Lagnoux, Sabine Mercier, Pierre Vallois

► **To cite this version:**

Agnes Lagnoux, Sabine Mercier, Pierre Vallois. Probability density function of the local score position. *Stochastic Processes and their Applications*, 2019, 129 (10), pp.3664-3689. 10.1016/j.spa.2018.10.008 . hal-01835781

HAL Id: hal-01835781

<https://hal.science/hal-01835781>

Submitted on 11 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

Probability density function of the local score position

Agnès Lagnoux¹, Sabine Mercier¹, Pierre Vallois^b

^a*Institut de Mathématiques de Toulouse, UMR5219, Université de Toulouse 2 Jean Jaurès, 5 allées Antonio Machado, 31058, Toulouse, Cedex 09, France,*

^b*Institut Elie Cartan, UMR7502 CNRS, INRIA-BIGS, Université de Lorraine, 54506 Vandoeuvre-lès-Nancy Cedex, France.*

Abstract

22
23
24
25
26
27
28
29
30
31

We calculate the probability density function of the local score position on complete excursions of a reflected Brownian motion. We use the trajectorial decomposition of the standard Brownian bridge to derive two different expressions of the density: the first one is based on a series and an integral while the second one is free off the series.

32
33
34
35
36

Keywords: Reflected Brownian motion ; Brownian bridge ; Brownian excursions ; local score ; sequence analysis.

2000 MSC: 60 G 17 ; 60 J 25 ; 60 J 65.

1. Introduction

37
38
39
40
41
42
43
44

This work is motivated by biological sequence analysis (e.g. DNA or proteins) largely developed since the 90's with the creation of databases. The local score is an usual tool to point out atypical segments of biological sequences. It has first been defined by Karlin and Altschul [9].

45
46
47
48

Let $(\varepsilon_i)_{i \geq 1}$ be a sequence of i.i.d. random variables, such that $\mathbb{E}[\varepsilon_i] = 0$ and $\text{Var}(\varepsilon_i) = 1$. The random walk $(S_n)_{n \geq 1}$ associated with $(\varepsilon_i)_{i \geq 1}$ is:

$$S_n := \sum_{i=1}^n \varepsilon_i, \quad n \geq 1, \quad S_0 = 0.$$

49
50
51
52
53
54
55
56
57

Email addresses: lagnoux@univ-tlse2.fr (Agnès Lagnoux),
mercier@univ-tlse2.fr (Sabine Mercier), pierre.vallois@univ-lorraine.fr
(Pierre Vallois)

58
59
60
61
62
63
64
65

Preprint submitted to Stochastic Processes and their Applications

January 15, 2018

The Lindley process $(U_n)_{n \geq 1}$ and the local score process $(H_n)_{n \geq 0}$ are respectively defined as:

$$U_n := S_n - \min_{0 \leq k \leq n} S_k, \quad n \geq 1, \quad (1.1)$$

$$H_n := \max_{0 \leq k \leq n} U_k = \max_{0 \leq i \leq j \leq n} (S_j - S_i), \quad n \geq 0. \quad (1.2)$$

This classical setting has been extended in many directions (see [10, 13, 8, 14]).

For biological applications, the distribution of H_n , when n is large, plays an important role. Here we do not study this question, the interested reader can consult [7, 6].

The Donsker theorem (see Section 2.10 in [1]) permits to define the local score in continuous time, see for instance Theorem 1 in [5] and [4]. The underlying process is the standard Brownian motion $(B(t); t \geq 0)$. It can be easily proved (see [4]) that the Lindley process $(U_k/\sqrt{n})_{0 \leq k \leq n}$ can be approximated by:

$$\widehat{U}(t) := B(t) - \inf_{0 \leq s \leq t} B(s), \quad 0 \leq t \leq 1. \quad (1.3)$$

Recall that $(\widehat{U}(t); t \geq 0)$ is distributed as the reflected Brownian motion $(U(t); t \geq 0)$ where

$$U(t) := |B(t)|, \quad t \geq 0. \quad (1.4)$$

According to (1.1) - (1.4), we define the local score in continuous time as:

$$\overline{U}(t) := \sup_{0 \leq s \leq t} U(s), \quad t \geq 0. \quad (1.5)$$

Let $f(t)$ be the unique time which achieves the maximum of U over $[0, t]$:

$$f(t) := \sup\{r \leq t; U(r) = \overline{U}(t)\}, \quad t \geq 0 \quad (1.6)$$

We say that the maximum $\overline{U}(t)$ occurs on a complete excursion if $f(t) \leq g(t)$, where

$$g(t) := \sup\{s \leq t; U(s) = 0\}, \quad t \geq 0. \quad (1.7)$$

We have calculated in [11] the probability of the event $\{f(t) \leq g(t)\}$. Here we deal with the local score U^* determined on complete excursions:

$$U^*(t) := \overline{U}(g(t)) = \sup_{0 \leq s \leq g(t)} U(s), \quad t \geq 0. \quad (1.8)$$

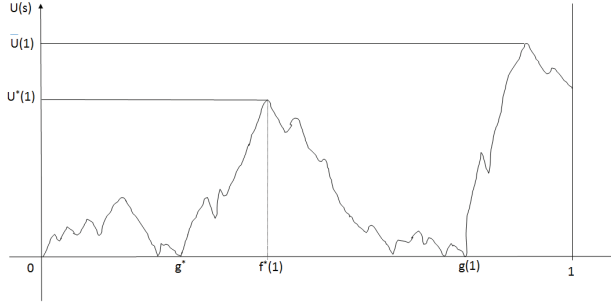


Figure 1: The r.v.'s $g(1)$, $\bar{U}(1)$, $U^*(1)$, $f^*(1)$ and g^*

Let $f^*(t)$ be the unique time which achieves the maximum of U over $[0, g(t)]$

$$f^*(t) := \sup\{r \leq g(t) ; U(r) = U^*(t)\}, \quad t \geq 0. \quad (1.9)$$

Now we are able to define the left end-point $g^*(t)$ of the excursion straddling $f^*(t)$

$$g^*(t) := g(f^*(t)) = \sup\{r \leq f^*(t) ; U(r) = 0\}, \quad t \geq 0. \quad (1.10)$$

In [4], we have calculated the probability density function of $(U^*(t), f^*(t) - g^*(t))$. Unfortunately, it is complicated but the density function of $U^*(t)$ is rather simple since it equals the sum of an explicit series. We focus here on the distribution of $g^*(t)$. This random time can be interpreted in the setting of local score.

Recall that the scaling property of the Brownian motion implies that $g^*(t)$ is distributed as $tg^*(1)$. For that reason, we only consider in the sequel $t = 1$. For simplicity, denote $g^*(1)$ by g^* . We have drawn a trajectory of $(U(t) ; 0 \leq t \leq 1)$ (see Figure 1) and indicated the variables introduced above.

2. The main result and the scheme of its proof

Before stating the main result of the paper, let us introduce

$$h(x) := \sum_{k \geq 1} (-1)^{k+1} \frac{k}{\cosh^2(kx)}. \quad (2.11)$$

Theorem 2.1. [Probability distribution function of g^*]

(i) The probability distribution function of g^* is given by

$$p_{g^*}(y) = \frac{1}{2\pi\sqrt{y(1-y)}} \int_0^{+\infty} \ln \left| 1 - \frac{\pi^2(1-y)}{4ys} \right| \frac{h(\sqrt{s})}{\sqrt{s}} ds, \quad 0 \leq y \leq 1. \quad (2.12)$$

(ii) The probability distribution function of g^* is also given by

$$p_{g^*}(y) = \frac{1}{\pi y} \int_0^{+\infty} \ln |\cot s| \frac{ds}{\cosh^2 \left(s \sqrt{\frac{1-y}{y}} \right)}, \quad 0 \leq y \leq 1. \quad (2.13)$$

This new result is in the stream of previous papers [4, 11, 12]. Among the three studies, two are theoretical [4, 11]. The last one is a review of asymptotic distributions of the local score in i.i.d. models and also contains illustrative simulations and statistical tests.

Let us explain the scheme of the proof of Theorem 2.1. Our approach is based on the identity (in distribution):

$$\left(\frac{1}{\sqrt{g(1)}} B(sg(1)); 0 \leq s \leq 1 \right) \stackrel{(d)}{=} (b(s); 0 \leq s \leq 1) \quad (2.14)$$

where $(b(s); 0 \leq s \leq 1)$ represents the standard Brownian bridge and

$$\left(\frac{1}{\sqrt{g(1)}} B(sg(1)); 0 \leq s \leq 1 \right) \text{ is independent of } g(1). \quad (2.15)$$

If we replace U by $|b|$ in the scheme (1.7)-(1.10) (resp. (1.7)-(1.8)), we get g^b (resp. $b^*(1)$) which have been plotted in Figure 2

Proposition 2.2. We have

$$g^* \stackrel{(d)}{=} g(1)g^b \quad (2.16)$$

where $g(1)$ and g^b are independent r.v.

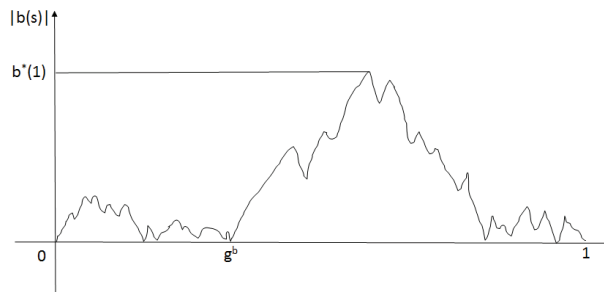


Figure 2: The r.v.'s g^b and $b^*(1)$

Proof of Proposition 2.2 Let $(b'(s); 0 \leq s \leq 1)$ be the process defined by

$$b'(s) := \frac{1}{\sqrt{g(1)}} B(sg(1)), \quad 0 \leq s \leq 1. \quad (2.17)$$

We easily deduce $g(1)g^{b'} = g^*$. Then (2.16) is a direct consequence of (2.14) and (2.15). \square

The probability density function of $g(1)$ is well known, see for instance Equation (2.22) in [4, Theorem 2.6]:

$$\mathbb{P}(g(1) \in dx) = \frac{1}{\pi \sqrt{x(1-x)}} \mathbb{1}_{[0,1]}(x) dx. \quad (2.18)$$

Hence, it remains to determine the distribution of g^b . One step in this direction comes from [15, Theorem 2]. Before stating the result, let us fix notation. Let $(L(t), t \geq 0)$ be the local time process at 0 related to the Brownian motion $(B(t), t \geq 0)$ and $(\tau_s, s \geq 0)$ be its right inverse. Let us consider the r.v. ξ distributed as $T_1(R)$ with

$$T_x(R) = \inf\{s \geq 0; R(s) = x\}, \quad x > 0. \quad (2.19)$$

$(R(s), s \geq 0)$, introduced in [4], stands for a 3-dimensional Bessel process started at 0 and will be extensively used in the sequel. The density p_ξ is explicitly known (see [2]) and is given by

$$p_\xi(u) = \frac{1}{\sqrt{2\pi}u^{3/2}} \sum_{k \in \mathbb{Z}} \left(-1 + \frac{(1+2k)^2}{u} \right) \exp \left(-\frac{(1+2k)^2}{2u} \right) \quad (2.20)$$

$$= \frac{d}{du} \left(\sum_{k \in \mathbb{Z}} (-1)^k \exp \left(-\frac{k^2 \pi^2 u}{2} \right) \right), \quad u > 0. \quad (2.21)$$

We always assume in the sequel that ξ and $(U(t))_{t \geq 0}$ are independent.

Proposition 2.3. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded Borel function. Then*

$$\mathbb{E}[f(g^b)] = \frac{\sqrt{2\pi}}{2} \int_{\mathbb{R}_+^2} \mathbb{E} \left[f \left(\frac{l^2 \tau_1}{l^2 \tau_1 + 4t} \right) \frac{1}{\sqrt{l^2 \tau_1 + 4t}} \mathbb{1}_{\{\bar{U}(\tau_1) < 1\}} \right] p_\xi(t) dl dt. \quad (2.22)$$

The proof of Proposition 2.3 is postponed in Section 3.1. Integrating with respect to l in (2.22) and making the change of variable $u = l^2 \tau_1 / (l^2 \tau_1 + 4t)$, considering that τ_1 and t are fixed, permits to obtain the density function of g^b .

Corollary 2.4. *[Probability distribution function of g^b]*

The probability distribution function of g^b is given by

$$p_{g^b}(u) = \frac{\sqrt{2\pi}}{4} \frac{1}{\sqrt{u}(1-u)} \int_0^{+\infty} \mathbb{E} \left[\frac{1}{\sqrt{\tau_1}} \mathbb{1}_{\{u(4t\bar{U}(\tau_1)^2 + \tau_1) < \tau_1\}} \right] p_\xi(t) dt \mathbb{1}_{[0,1]}(u). \quad (2.23)$$

Identities (2.12) and (2.13) follow from Proposition 2.2, Corollary 2.4 and numerous calculations which are developed in Sections 3.2, 3.3 and 3.4. The proof of Proposition 2.3 and Theorem 2.1 are derived in a series of steps. In order to keep a certain fluidity, all the proofs which are not immediate have been given in the last Section 3.4.

3. Proofs

We keep the notation introduced in Section 2.

3.1. Proof of Proposition 2.3

1) By definition, g^b is defined via steps (1.7)-(1.10) where we substitute $|b|$ for U . Theorem 2 [15] is convenient for our purpose since it expresses the law of $(b(s); 0 \leq s \leq 1)$ conditionally on b^* , with two independent pieces of Brownian trajectories. Let us introduce

$$br^*(u) = \begin{cases} \frac{1}{\sqrt{\sigma_1 + \hat{\sigma}_1}} B(u(\sigma_1 + \hat{\sigma}_1)), & 0 \leq u \leq \frac{\sigma_1}{\sigma_1 + \hat{\sigma}_1} \\ \frac{1}{\sqrt{\sigma_1 + \hat{\sigma}_1}} \hat{B}((1-u)(\sigma_1 + \hat{\sigma}_1)), & \frac{\sigma_1}{\sigma_1 + \hat{\sigma}_1} \leq u \leq 1 \end{cases} \quad (3.24)$$

where B and \hat{B} are two independent standard Brownian motions and

$$\sigma_1 = \inf\{t \geq 0; B(t) = 1\} \quad \text{and} \quad \hat{\sigma}_1 = \inf\{t \geq 0; \hat{B}(t) = 1\}.$$

$(br^*(u); 0 \leq u \leq 1)$ is the concatenation of $(B(u); 0 \leq u \leq \sigma_1)$ and $(\hat{B}(t); 0 \leq t \leq \hat{\sigma}_1)$ with the scaling in space (respectively time) $1/\sqrt{\sigma_1 + \hat{\sigma}_1}$ (resp. $1/(\sigma_1 + \hat{\sigma}_1)$).

Lemma 3.1. [Theorem 2 in [15]] For every non-negative measurable function F defined on the path space $C([0, 1])$,

$$\mathbb{E}[F(b(u); 0 \leq u \leq 1)] = \sqrt{2\pi} \mathbb{E}[F(br^*(u); 0 \leq u \leq 1)M^*] \quad (3.25)$$

where

$$M^* := \sup_{0 \leq u \leq 1} br^*(u) = \frac{1}{\sqrt{\sigma_1 + \hat{\sigma}_1}}. \quad (3.26)$$

We use the structure of the trajectory of br^* and (3.25) to give a first expression of the distribution of g^b in terms of Brownian random variables.

Corollary 3.2. Let f be a bounded Borel function, then

$$\mathbb{E}[f(g^b)] = 2\sqrt{2\pi} \mathbb{E} \left[f \left(\frac{g(\sigma_1)}{\sigma_1 + \hat{\sigma}_1} \right) \frac{1}{\sqrt{\sigma_1 + \hat{\sigma}_1}} \mathbb{1}_{\{\underline{B}(\sigma_1) \wedge \underline{\hat{B}}(\hat{\sigma}_1) > -1\}} \right] \quad (3.27)$$

where $\underline{B}(\sigma_1) := \inf_{0 \leq u \leq \sigma_1} B(u)$ and $\underline{\hat{B}}(\hat{\sigma}_1) := \inf_{0 \leq u \leq \hat{\sigma}_1} \hat{B}(u)$.

Proof of Corollary 3.2 First, notice that $b^*(1) = \sup_{0 \leq u \leq 1} |b(u)| = \bar{b}(1) \vee \underline{b}(1)$, where $\bar{b}(1) := \sup_{0 \leq u \leq 1} b(u)$ and $\underline{b}(1) := -\inf_{0 \leq u \leq 1} b(u)$. Moreover, $b^*(1) = \bar{b}(1) \Leftrightarrow \bar{b}(1) > \underline{b}(1)$ and on that set, $b^*(1) = \bar{b}(1)$. Since $(-b(u); 0 \leq$

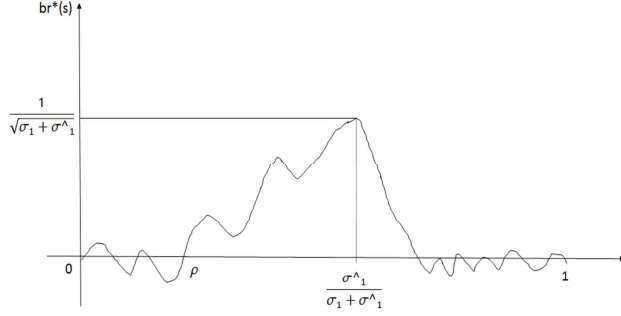


Figure 3: Notation $br^*(1)$ and ρ

$u \leq 1$) is also a Brownian bridge, we apply Lemma 3.1 to $F(b(u); 0 \leq u \leq 1) = f(g^b)$ to get:

$$\begin{aligned} \mathbb{E}[f(g^b)] &= 2\mathbb{E}[f(g^b)\mathbb{1}_{\{\bar{b}(1) > \underline{b}(1)\}}] \\ &= 2\sqrt{2\pi}\mathbb{E}\left[\frac{f(\rho)}{\sqrt{\sigma_1 + \hat{\sigma}_1}}\mathbb{1}_{\left\{\sup_{0 \leq u \leq 1} br^*(u) > -\inf_{0 \leq u \leq 1} br^*(u)\right\}}\right], \end{aligned}$$

where ρ is the starting point of the highest excursion of $(br^*(s); 0 \leq s \leq 1)$. See Figure 3 for more precisions. We remark that $(\sigma_1 + \hat{\sigma}_1)\rho = g(\sigma_1)$ (see Figure 4) and

$$\begin{aligned} \left\{\sup_{0 \leq u \leq 1} br^*(u) > -\inf_{0 \leq u \leq 1} br^*(u)\right\} &= \left\{\inf_{0 \leq u \leq \sigma_1 + \hat{\sigma}_1} X(u) > -1\right\} \\ &= \left\{\inf_{0 \leq u \leq \sigma_1} B(u) \wedge \inf_{0 \leq u \leq \hat{\sigma}_1} \hat{B}(u) > -1\right\}. \end{aligned}$$

Identity (3.27) follows immediately. \square

2) Remark that the expectation in the right hand side of (3.27) involves the law of

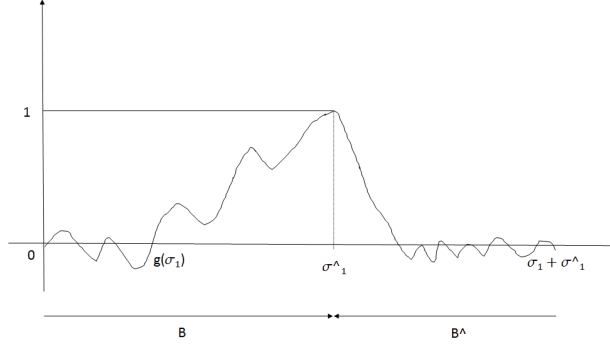


Figure 4: Scaling

($g(\sigma_1), \sigma_1, \underline{B}(\sigma_1), \hat{\sigma}_1, \hat{\underline{B}}(\hat{\sigma}_1)$). Since B and \hat{B} are independent,

$$\text{Law} \left(g(\sigma_1), \sigma_1, \underline{B}(\sigma_1), \hat{\sigma}_1, \hat{\underline{B}}(\hat{\sigma}_1) \right) = \text{Law} \left(g(\sigma_1), \sigma_1, \underline{B}(\sigma_1) \right) \otimes \text{Law} \left(\hat{\sigma}_1, \hat{\underline{B}}(\hat{\sigma}_1) \right). \quad (3.28)$$

First we determine the law of $(\hat{\sigma}_1, \hat{\underline{B}}(\hat{\sigma}_1))$, see Lemma 3.3 below. Second we consider more generally the law of $(g(\sigma_1), \sigma_1, \underline{B}(\sigma_1))$, see Lemma 3.5.

Lemma 3.3. *Let $a > 0$ and $\varphi : [0, +\infty[\rightarrow [0, +\infty[$. Then*

$$\mathbb{E}[\varphi(\sigma_1) \mathbb{1}_{\{\underline{B}(\sigma_1) > -a\}}] = \int_0^{+\infty} \varphi(t) s s_t(a, a+1) dt$$

where

$$s s_t(x, z) = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \frac{z - x + 2kz}{t^{3/2}} \exp\{- (z - x + 2kz)^2 / (2t)\}. \quad (3.29)$$

Proof of Lemma 3.3 We have

$$\mathbb{E}[\varphi(\sigma_1) \mathbb{1}_{\{\underline{B}(\sigma_1) > -a\}}] = \mathbb{E}[\varphi(\sigma_1) | \underline{B}(\sigma_1) > -a] \mathbb{P}(\underline{B}(\sigma_1) > -a).$$

By definition, $\mathbb{P}(\underline{B}(\sigma_1) > -a) = \mathbb{P}(\inf_{t \in [0, \sigma_1]} B(t) > -a) = \mathbb{P}(\sigma_1 < \sigma_{-a}) = a/(a+1)$

and conditionally on $\{\underline{B}(\sigma_1) > -a\}$, $(a + B(t), 0 \leq t \leq \sigma_1) \stackrel{(d)}{=} (R_a(t), 0 \leq t \leq H_{a+1})$ where $(R_a(t), t \geq 0)$ is a 3-dimensional Bessel process started at a and

$$H_z = \inf\{t \geq 0, R_a(t) = z\}.$$

Assume that $a < z$. By Formula 2.0.2 p63 in [3],

$$\mathbb{P}(H_z \in dt) = \frac{z}{a} ss_t(a, z) dt. \quad (3.30)$$

□

Using (3.28), Corollary 3.2 and Lemma 3.3, we deduce:

Corollary 3.4. *Let f be a bounded Borel function, then*

$$\mathbb{E}[f(g^b)] = 2\sqrt{2\pi} \int_0^{+\infty} \mathbb{E} \left[\mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}} f \left(\frac{g(\sigma_1)}{\sigma_1 + t} \right) \frac{1}{\sqrt{\sigma_1 + t}} \right] ss_t(1, 2) dt. \quad (3.31)$$

Now we deal with the distribution of $(g(\sigma_1), \sigma_1)$ conditionally on $\{\underline{B}(\sigma_1) > -1\}$.

Lemma 3.5. *Let φ be a bounded Borel function. Then,*

$$\mathbb{E}[\varphi(g(\sigma_1), \sigma_1) \mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}}] = \frac{1}{2} \int_{\mathbb{R}_+^2} \mathbb{E} \left[\varphi(l^2 \tau_1, l^2 \tau_1 + t) \mathbb{1}_{\{\bar{u}(\tau_1) < 1\}} \right] p_\xi(t) dl dt \quad (3.32)$$

Proof of Lemma 3.5 From Proposition 4 in [17], $\sigma_1 - g(\sigma_1)$ is independent of $(g(\sigma_1), \underline{B}_{\sigma_1})$ and distributed as ξ (see Section 2). Moreover, $\underline{B}_{\sigma_1} = \underline{B}_{g(\sigma_1)}$. Hence

$$\mathbb{E}[\varphi(g(\sigma_1), \sigma_1) \mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}}] = \int_0^{+\infty} \mathbb{E}[\varphi(g(\sigma_1), g(\sigma_1) + t) \mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}}] p_\xi(t) dt.$$

Recall that $(L(t); t \geq 0)$ is the local time at 0 of $(B(t); t \geq 0)$. Using several times Proposition 4 in [17] we get:

$$\mathbb{E}[\varphi(g(\sigma_1), \sigma_1) \mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}}]$$

$$\begin{aligned}
&= \int_0^{+\infty} p_\xi(t) \int_0^{+\infty} \frac{1}{2} e^{-l/2} \mathbb{E}[\varphi(g(\sigma_1), g(\sigma_1) + t) \mathbb{1}_{\{\underline{B}(\sigma_1) > -1\}} | L(\sigma_1) = l] dl dt \\
&= \frac{1}{2} \int_{\mathbb{R}_+^2} \mathbb{E}[\varphi(\tau_l, \tau_l + t) \mathbb{1}_{\{\underline{B}(\tau_l) > -1\}} | \overline{B}(\tau_l) < 1] \mathbb{P}(\overline{B}(\tau_l) < 1) p_\xi(t) dl dt \\
&= \frac{1}{2} \int_{\mathbb{R}_+^2} \mathbb{E}[\varphi(\tau_l, \tau_l + t) \mathbb{1}_{\{\underline{B}(\tau_l) > -1, \overline{B}(\tau_l) < 1\}}] p_\xi(t) dl dt.
\end{aligned}$$

Using $U = |B|$ and the scaling identity $(\overline{U}(\tau_l), \tau_l) \stackrel{(d)}{=} (l\overline{U}(\tau_1), l^2\tau_1)$ (see Formula (4.2) in [4]) ends the proof of Lemma 3.5. \square

Remark 3.6. *The law of $(\tau_1, \overline{U}(\tau_1))$ is rather complicated, see Proposition 2.2 in [4]. Lemma 3.5 does not permit to have easily an explicit formula of the density function of $g(\sigma_1)$ conditionally on $\{\underline{B}_{\sigma_1} > -1\}$ and therefore to recover Lemma 3.3.*

Now we are able to end the proof of Proposition 2.3. Using Corollary 3.4 and Lemma 3.5, one gets

$$\begin{aligned}
\mathbb{E}[f(g^b)] &= \sqrt{2\pi} \int_{\mathbb{R}_+^3} \mathbb{E} \left[f \left(\frac{l^2\tau_1}{l^2\tau_1 + t_1 + t_2} \right) \frac{1}{\sqrt{l^2\tau_1 + t_1 + t_2}} \mathbb{1}_{\{l\overline{U}(\tau_1) < 1\}} \right] \\
&\quad \cdot p_\xi(t_1) ss_{t_2}(1, 2) dl dt_1 dt_2 \\
&= \frac{\sqrt{2\pi}}{2} \int_{\mathbb{R}_+^2} \mathbb{E} \left[f \left(\frac{l^2\tau_1}{l^2\tau_1 + 4t} \right) \frac{1}{\sqrt{l^2\tau_1 + 4t}} \mathbb{1}_{\{l\overline{U}(\tau_1) < 1\}} \right] p_\xi(t) dl dt
\end{aligned}$$

because $p_\xi * 2ss.(1, 2)$ is the probability distribution function of H_2 when $R_0 = 0$ and $H_2 \stackrel{(d)}{=} 4H_1 \stackrel{(d)}{=} 4\xi$ by the scaling property of the Bessel process. \square

3.2. Proof of (2.12) in Theorem 2.1

3.2.1. A first formula of the probability distribution function of g^*

Our starting point is obviously Proposition 2.2. Note that the density function of $g(1)$ is explicit and given by (2.18) and (2.23) shows that the one of g^b is expressed via the unknown quantity:

$$I(u) := \int_{\mathbb{R}_+} \mathbb{E} \left[\frac{1}{\sqrt{\tau_1}} \mathbb{1}_{\{u(4t\overline{U}(\tau_1)^2 + \tau_1) < \tau_1\}} \right] p_\xi(t) dt, \quad 0 < u < 1. \quad (3.33)$$

The main result of this subsection is Corollary 3.9.

Lemma 3.7. For any $0 < y < 1$, one has

$$p_{g^*}(y) = -\frac{1}{2\sqrt{2\pi}} \frac{1}{\sqrt{y}} \lim_{\varepsilon \rightarrow 0} \int_y^{1-\varepsilon} \left(\int_y^s \frac{du}{(1-u)\sqrt{u(u-y)}} \right) dI(s). \quad (3.34)$$

Proof of Lemma 3.7 Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded Borel function. Using Proposition 2.2, Corollary 2.4 and (2.18), we obtain:

$$\begin{aligned} \mathbb{E}[f(g^*)] &= \frac{1}{\pi} \int_0^1 \frac{\mathbb{E}[f(xg^b)]}{\sqrt{x(1-x)}} dx \\ &= \frac{1}{2\sqrt{2\pi}} \iint_0^1 \frac{f(y)}{\sqrt{y(u-y)}} \frac{I(u)}{\sqrt{u(1-u)}} \mathbb{1}_{\{0 \leq y \leq u\}} dy du \end{aligned} \quad (3.35)$$

where the last equality comes from the change of variable $ux = y$ for a fixed u . Hence, using that $I(1) = 0$ and I is decreasing:

$$\begin{aligned} p_{g^*}(y) &= \frac{1}{2\sqrt{2\pi}} \frac{1}{\sqrt{y}} \int_y^1 \frac{I(u)}{\sqrt{u(u-y)(1-u)}} du \\ &= -\frac{1}{2\sqrt{2\pi}} \frac{1}{\sqrt{y}} \int_y^1 \frac{1}{\sqrt{u(u-y)(1-u)}} \left(\int_u^1 dI(s) \right) du \\ &= -\frac{1}{2\sqrt{2\pi}} \frac{1}{\sqrt{y}} \lim_{\varepsilon \rightarrow 0} \int_y^{1-\varepsilon} \left(\int_y^s \frac{du}{\sqrt{u(u-y)(1-u)}} \right) dI(s). \quad \square \end{aligned}$$

As shows (3.35) there is a singularity at $u = 1$. This explains why we introduce the cutoff ε . Now our strategy is to compute I (Lemma 3.8) and then its derivative. This leads to Corollary 3.9, which is the first main step of the proof of (2.12).

Lemma 3.8. We get

$$I(u) = \sqrt{\frac{2}{\pi}} + 2 \sum_{k \geq 1} (-1)^k I_1 \left(\frac{k^2 \pi^2}{8} \frac{1-u}{u} \right), \quad 0 < u < 1 \quad (3.36)$$

where

$$I_1(v) := \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{s} \cosh^2(\sqrt{s+2v})} ds. \quad (3.37)$$

Proof of Lemma 3.8 1) Let h_1 be the real-valued function defined by

$$h_1(u) := 2 \sum_{k \geq 1} (-1)^k \exp \left\{ -\frac{k^2 \pi^2 u}{2} \right\}, \quad u > 0. \quad (3.38)$$

By (2.21), $p_\xi = h'_1$, then for any $t_1 > 0$,

$$\int_0^{t_1} p_\xi(t) dt = \int_0^{+\infty} p_\xi(t) dt - \int_{t_1}^{+\infty} h'_1(t) dt = 1 + h_1(t_1).$$

Since $\tau_1 \stackrel{(d)}{=} 1/N^2$ with $N \sim \mathcal{N}(0, 1)$, the above identity and (3.38) imply:

$$\begin{aligned} I(u) &= \mathbb{E} \left[\frac{1}{\sqrt{\tau_1}} \left(1 + h_1 \left(\frac{1-u}{u} \frac{\tau_1}{4\overline{U}(\tau_1)^2} \right) \right) \right] \\ &= \sqrt{\frac{2}{\pi}} + 2 \sum_{k \geq 1} (-1)^k I_1 \left(\frac{k^2 \pi^2}{8} \frac{1-u}{u} \right) \end{aligned}$$

where

$$I_1(v) = \mathbb{E} \left[\frac{1}{\sqrt{\tau_1}} \exp \left\{ -v \frac{\tau_1}{\overline{U}(\tau_1)^2} \right\} \right], \quad v > 0. \quad (3.39)$$

2) We calculate I_1 . Using the identity

$$\int_0^{+\infty} \frac{1}{\sqrt{t}} e^{-ta} dt = \sqrt{\frac{\pi}{a}},$$

with $a = \tau_1$ and inverting \int and \mathbb{E} in (3.39) leads to:

$$\begin{aligned} I_1(v) &= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \mathbb{E} \left[\exp \left\{ -\tau_1 \left(t + \frac{v}{\overline{U}(\tau_1)^2} \right) \right\} \right] \frac{dt}{\sqrt{t}} \\ &= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}_+^2} \mathbb{E} \left[\exp \left\{ -\tau_1 \left(t + \frac{v}{x^2} \right) \right\} \mid \overline{U}(\tau_1) = x \right] \frac{1}{x^2} e^{-1/x} dx \frac{dt}{\sqrt{t}}, \end{aligned}$$

since $1/\overline{U}(\tau_1)$ is exponentially distributed (see [16, Theorem 1]). We claim:

$$\mathbb{E} \left[e^{-\mu \tau_1} \mid \overline{U}(\tau_1) = x \right] \frac{e^{-1/x}}{x^2} = \frac{2\mu}{\sinh^2(x\sqrt{2\mu})} \exp \left\{ -\sqrt{2\mu} \coth(x\sqrt{2\mu}) \right\}, \quad \mu > 0. \quad (3.40)$$

Indeed, by Proposition 2.2 and Formula (4.13) in [4],

$$\mathbb{E} \left[e^{-\mu\tau_1} \mathbb{1}_{\{\bar{U}(\tau_1) < x\}} \right] = \exp \left\{ -\sqrt{2\mu} \coth(x\sqrt{2\mu}) \right\} \quad (3.41)$$

$$= \int_0^x \mathbb{E} \left[e^{-\mu\tau_1} | \bar{U}(\tau_1) = y \right] \frac{1}{y^2} e^{-1/y} dy, \quad (3.42)$$

$$(3.43)$$

and a derivative with respect to x leads to (3.40).

Using (3.40) and the change of variable $u = 2(tx^2 + v)$ for a fixed x ,

$$\begin{aligned} I_1(v) &= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}_+^2} \frac{2(tx^2 + v)}{\sinh^2 \left(\sqrt{2(tx^2 + v)} \right)} \\ &\quad \cdot \exp \left\{ -\sqrt{2 \left(t + \frac{v}{x^2} \right)} \coth \left(\sqrt{2(tx^2 + v)} \right) \right\} \frac{dxdt}{x^2\sqrt{t}} \\ &= \frac{1}{2\sqrt{\pi}} \int_{\mathbb{R}_+^2} \frac{u}{x^3 \sqrt{\frac{u}{2} - v} \sinh^2(\sqrt{u})} \exp \left\{ -\frac{\sqrt{u}}{x} \coth(\sqrt{u}) \right\} \mathbb{1}_{\{u > 2v\}} dudx \\ &= \frac{1}{2\sqrt{\pi}} \int_{2v}^{+\infty} \frac{1}{\sqrt{\frac{u}{2} - v} \cosh^2(\sqrt{u})} \left(\int_0^{+\infty} we^{-w} dw \right) du \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{s} \cosh^2(\sqrt{s + 2v})} ds \end{aligned}$$

where we make the change of variable $w = \sqrt{u} \coth(\sqrt{u})/x$ for a fixed u . \square

Corollary 3.9. *We have*

$$p_{g^*}(y) = \frac{\pi}{8\sqrt{y}} \lim_{\varepsilon \rightarrow 0} \sum_{k \geq 1} (-1)^{k+1} k^2 \delta_k(y, \varepsilon), \quad 0 < y < 1, \quad (3.44)$$

where

$$\begin{aligned} \delta_k(y, \varepsilon) &:= \int_0^{+\infty} \frac{1}{\sqrt{v}} \int_y^{1-\varepsilon} \frac{1}{s^2} \frac{1}{\sqrt{\left(v + \frac{k^2\pi^2}{4} \frac{1-s}{s}\right)}} \frac{\sinh}{\cosh^3} \left(\sqrt{v + \frac{k^2\pi^2}{4} \frac{1-s}{s}} \right) \\ &\quad \times \left(\int_y^s \frac{1}{(1-u)\sqrt{u(u-y)}} du \right) dsdv. \end{aligned} \quad (3.45)$$

Proof of Corollary 3.9 We take the derivative in (3.36) and (3.37) and we get

$$\begin{aligned}
I'(s) &= -\frac{\pi^2}{4s^2} \sum_{k \geq 1} (-1)^k k^2 I'_1 \left(\frac{k^2 \pi^2 (1-s)}{8s} \right) \\
&= \frac{\pi^2}{4s^2} \sqrt{\frac{2}{\pi}} \sum_{k \geq 1} (-1)^k k^2 \\
&\quad \cdot \int_0^{+\infty} \frac{1}{\sqrt{v \left(v + \frac{k^2 \pi^2 (1-s)}{4s} \right)}} \frac{\sinh}{\cosh^3} \left(\sqrt{v + \frac{k^2 \pi^2 (1-s)}{4s}} \right) dv \quad (3.46)
\end{aligned}$$

that plugged into (3.34) leads to the required result. \square

*3.2.2. A second formula of the probability distribution function of g^**

From (3.45), we see that $\delta_k(y, \varepsilon)$ is a triple integral. We reduce it to a simple integral see Lemma 3.10 below. Then (3.44) implies that $p_{g^*}(y)$ is a limit of a single integral (see Corollary 3.11). The limit can be calculated, cf. Lemma 3.12, that proves (3.44).

Lemma 3.10. *Letting $\varepsilon' = \varepsilon/1 - \varepsilon$, one has*

$$\delta_k(y, \varepsilon) = \frac{4}{k^2 \pi^2} \frac{1}{\sqrt{1-y}} \int_0^{+\infty} \left(\frac{\rho_1(\varepsilon')}{\sqrt{s}} - \frac{\rho_2(s, \varepsilon')}{\sqrt{s + \frac{\pi^2}{4} \varepsilon'}} \right) \frac{k ds}{\cosh^2 \left(k \sqrt{s + \frac{\pi^2}{4} \varepsilon'} \right)} \quad (3.47)$$

where

$$\begin{aligned}
\rho_1(\varepsilon') &:= \ln \left(\frac{\sqrt{1-y} + \sqrt{1-y(1+\varepsilon')}}{\sqrt{1-y} - \sqrt{1-y(1+\varepsilon')}} \right) \\
\text{and } \rho_2(s, \varepsilon') &:= \ln \left(\frac{u_1(s, \varepsilon') + u_2(s, \varepsilon')}{|u_1(s, \varepsilon') - u_2(s, \varepsilon')|} \right) \quad (3.48)
\end{aligned}$$

with

$$u_1(s, \varepsilon') := \sqrt{\frac{4s}{\pi^2(1-y-y\varepsilon')}} \quad \text{and} \quad u_2(s, \varepsilon') := \sqrt{\frac{\frac{4s}{\pi^2} + \varepsilon'}{1-y}}. \quad (3.49)$$

Proof of Lemma 3.10 In (3.45), we directly integrate with respect to s and then we make the change of variable $w = (1 - u)/u$:

$$\begin{aligned}
& \delta_k(y, \varepsilon) \\
&= \frac{4}{k^2\pi^2} \int_{\mathbb{R}_+^2} \frac{\mathbb{1}_{\{y \leq u \leq 1-\varepsilon\}}}{(1-u)\sqrt{uv(u-y)}} \\
&\quad \cdot \left(\frac{1}{\cosh^2\left(\sqrt{v + \frac{k^2\pi^2}{4}\varepsilon'}\right)} - \frac{1}{\cosh^2\left(\sqrt{v + \frac{k^2\pi^2}{4}\frac{1-u}{u}}\right)} \right) dudv \\
&= \frac{4}{k^2\pi^2} \int_{\varepsilon'}^{\frac{1-y}{y}} \frac{dw}{w\sqrt{1-y(1+w)}} \\
&\quad \cdot \int_{\mathbb{R}_+} \left(\frac{1}{\cosh^2\left(\sqrt{v + \frac{k^2\pi^2}{4}\varepsilon'}\right)} - \frac{1}{\cosh^2\left(\sqrt{v + \frac{k^2\pi^2}{4}w}\right)} \right) \frac{dv}{\sqrt{v}} \\
&= \frac{4}{k^2\pi^2} (q_1 - q_2)
\end{aligned}$$

where

$$\begin{aligned}
q_1 &:= \left(\int_{\varepsilon'}^{\frac{1-y}{y}} \frac{dw}{w\sqrt{1-y(1+w)}} \right) \left(\int_0^{+\infty} \frac{dv}{\sqrt{v} \cosh^2\left(\sqrt{v + \frac{k^2\pi^2}{4}\varepsilon'}\right)} \right) \\
q_2 &:= \int_{\varepsilon'}^{\frac{1-y}{y}} \frac{dw}{w\sqrt{1-y(1+w)}} \left(\int_{\frac{k^2\pi^2}{4}w}^{+\infty} \frac{1}{\sqrt{t - \frac{k^2\pi^2}{4}w}} \frac{dt}{\cosh^2(\sqrt{t})} \right) \\
&= \int_{\frac{k^2\pi^2}{4}\varepsilon'}^{+\infty} \frac{q_3(t)}{\cosh^2(\sqrt{t})} dt \tag{3.50}
\end{aligned}$$

with

$$q_3(t) := \frac{2}{k\pi} \int_{\varepsilon'}^{t_0} \frac{1}{w\sqrt{1-y(1+w)}} \frac{1}{\sqrt{t_0-w}} \mathbb{1}_{\{w < \frac{1-y}{y}\}} dw \quad \text{and} \quad t_0 = \frac{4t}{k^2\pi}.$$

1
2
3
4
5
6
7
8
9
10
11 1) Observe that

$$12 \quad w \mapsto -\frac{1}{\sqrt{1-y}} \ln \left(\frac{\sqrt{1-y} + \sqrt{1-y(1+w)}}{\sqrt{1-y} - \sqrt{1-y(1+w)}} \right)$$

13
14
15
16
17 is a primitive of $1/w\sqrt{1-y(1+w)}$, therefore

$$18 \quad \int_{\varepsilon'}^{\frac{1-y}{y}} \frac{dw}{w\sqrt{1-y(1+w)}} = \frac{1}{\sqrt{1-y}} \ln \left(\frac{\sqrt{1-y} + \sqrt{1-y(1+\varepsilon')}}{\sqrt{1-y} - \sqrt{1-y(1+\varepsilon')}} \right)$$

$$19 \quad = \frac{1}{\sqrt{1-y}} \rho_1(\varepsilon')$$

20
21
22
23
24
25
26
27
28 and

$$29 \quad q_1 = \frac{1}{\sqrt{1-y}} \rho_1(\varepsilon') \int_0^{+\infty} \frac{k ds}{\sqrt{s} \cosh^2 \left(k \sqrt{s + \frac{\pi^2}{4} \varepsilon'} \right)} \quad (3.51)$$

30
31
32
33
34
35 after the change of variable $v = k^2 s$.

36
37 **2)** Now we simplify $q_3(t)$. Let $u = \sqrt{(t_0 - w)/(1 - y(1 + w))}$. From this
38 change of variable and we get:

- 39
40
41 • if $t_0 > (1 - y)/y$,

$$42 \quad q_3(t) = \frac{4}{k\pi} \int_{u_0}^{+\infty} \frac{du}{u^2(1-y) - t_0} = \frac{1}{\sqrt{t(1-y)}} \ln \left(\frac{u_0 + \sqrt{\frac{t_0}{1-y}}}{u_0 - \sqrt{\frac{t_0}{1-y}}} \right),$$

- 43
44
45
46
47
48
49 • if $t_0 < (1 - y)/y$,

$$50 \quad q_3(t) = \frac{4}{k\pi} \int_{u_0}^0 \frac{du}{u^2(1-y) - t_0} = \frac{1}{\sqrt{t(1-y)}} \ln \left(\frac{u_0 + \sqrt{\frac{t_0}{1-y}}}{\sqrt{\frac{t_0}{1-y}} - u_0} \right)$$

51
52
53
54
55
56 where $u_0 = \sqrt{(t_0 - \varepsilon')/(1 - y(1 + \varepsilon'))}$.

We make the change of variable $t = k^2 (s + \pi^2 \varepsilon' / 4)$ in (3.50), then $\sqrt{t_0 / (1 - y)} = u_2(s, \varepsilon')$, $u_0 = u_1(s, \varepsilon')$ and

$$q_3(t) = \frac{1}{k} \frac{1}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} \rho_2(s, \varepsilon').$$

Finally, (3.47) is a consequence of (3.51), (3.50) and the above identity. \square

Using identities (3.44) and (3.47), the definition (2.11) of h in and interverting (legally) the integral and the sum gives the following corollary because the cutoff ε' permits to intervert the integral and the sum. Finally, we can write $p_{g^*}(y)$ as a limit of an integral of a function which is expressed as a series.

Corollary 3.11. *We have*

$$p_{g^*}(y) = \frac{1}{2\pi\sqrt{y(1-y)}} \lim_{\varepsilon' \rightarrow 0} \int_0^{+\infty} \left(\frac{\rho_1(\varepsilon')}{\sqrt{s}} - \frac{\rho_2(s, \varepsilon')}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} \right) h \left(\sqrt{s + \frac{\pi^2 \varepsilon'}{4}} \right) ds \quad (3.52)$$

where h , ρ_1 and ρ_2 have respectively been defined in (2.11) and (3.48).

Here we compute the limit of the term $\rho_1(\varepsilon')/\sqrt{s} - \rho_2(s, \varepsilon')/\sqrt{s + \pi^2 \varepsilon' / 4}$ appearing in the previous integral and we verify that the dominated convergence theorem may be applied.

Lemma 3.12. *For any $s > 0$,*

$$\lim_{\varepsilon' \rightarrow 0} \left(\frac{\rho_1(\varepsilon')}{\sqrt{s}} - \frac{\rho_2(s, \varepsilon')}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} \right) h \left(\sqrt{s + \frac{\pi^2 \varepsilon'}{4}} \right) = \ln \left| 1 - \frac{\pi^2(1-y)}{4ys} \right| \frac{h(\sqrt{s})}{\sqrt{s}}. \quad (3.53)$$

Proof of Lemma 3.12 We use the following decomposition

$$\frac{\rho_1(\varepsilon')}{\sqrt{s}} - \frac{\rho_2(s, \varepsilon')}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} = \frac{\rho_3(\varepsilon')}{\sqrt{s}} - \frac{\rho_4(s, \varepsilon')}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} + \frac{\rho_5(s, \varepsilon')}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} + \ln \varepsilon' \left(\frac{1}{\sqrt{s + \frac{\pi^2 \varepsilon'}{4}}} - \frac{1}{\sqrt{s}} \right) \quad (3.54)$$

where

$$\rho_3(\varepsilon') = 2 \ln \left(\sqrt{1-y} + \sqrt{1-y(1+\varepsilon')} \right) - \ln(y), \quad (3.55)$$

$$\begin{aligned} \rho_4(s, \varepsilon') &= 2 \ln(u_1(s, \varepsilon') + u_2(s, \varepsilon')) + \ln(\pi^2(1-y)) \\ &\quad + \ln(1-y(1+\varepsilon')) - \ln(4y), \end{aligned} \quad (3.56)$$

$$\rho_5(s, \varepsilon') = \ln \left| s - \frac{\pi^2}{4} \frac{1-y}{y} + \frac{\pi^2}{4} \varepsilon' \right|. \quad (3.57)$$

It is easy to deduce from (3.49) and (2.11):

$$\begin{aligned} \lim_{\varepsilon' \rightarrow 0} \rho_3(\varepsilon') &= 2 \ln 2 + \ln(1-y) - \ln(y), \\ \lim_{\varepsilon' \rightarrow 0} \rho_4(s, \varepsilon') / \sqrt{s + \frac{\pi^2}{4} \varepsilon'} &= \ln(4s(1-y)/y) / \sqrt{s}, \\ \lim_{\varepsilon' \rightarrow 0} \rho_5(s, \varepsilon') &= \ln |s - \pi^2(1-y)/4y|, \\ \lim_{\varepsilon' \rightarrow 0} h \left(\sqrt{s + \frac{\pi^2}{4} \varepsilon'} \right) &= h(\sqrt{s}). \end{aligned}$$

Then (3.53) follows from (3.54). \square

3.2.3. Proof of (2.12)

We claim that we can invert in (3.52) the limit ($\varepsilon' \rightarrow 0$) and the integral, using the dominated convergence theorem. Then (2.12) is a direct consequence of Lemma 3.12. Indeed, there exists a generic constant $C > 0$ (that may change from one line to another) such that we have the following inequalities:

$$\begin{aligned} \frac{|\rho_3(\varepsilon')|}{\sqrt{s}} &\leq \frac{C}{\sqrt{s}}, \quad \forall s > 0, \quad \forall \varepsilon' \in (0, 1), \\ \frac{|\rho_4(s, \varepsilon')|}{\sqrt{s + \frac{\pi^2}{4} \varepsilon'}} &\leq \frac{C}{\sqrt{s}} (1 + \ln(1+s) + |\ln s|), \quad \forall s > 0, \quad \forall \varepsilon' \in (0, 1], \\ \frac{|\rho_5(s, \varepsilon')|}{\sqrt{s + \frac{\pi^2}{4} \varepsilon'}} &\leq \frac{C}{\sqrt{s}} \left(1 + \ln(1+s) + \left| \ln \left| s - \frac{\pi^2(1-y)}{4y} \right| \right| \right), \quad s \in I_{\varepsilon'}^c \end{aligned}$$

where

$$I_{\varepsilon'}^c := \left] 0, \frac{\pi^2(1-y)}{4y} - \pi^2\varepsilon' \left[\cup \left[\frac{\pi^2(1-y)}{4y}, +\infty \left[.$$

It remains to investigate the limit of the integral when $s \in I_{\varepsilon'}$. For $y > 0$ fixed and ε' small enough,

$$\frac{1}{\sqrt{s + \frac{\pi^2}{4}\varepsilon'}} \leq C, \quad h\left(\sqrt{s + \frac{\pi^2}{4}\varepsilon'}\right) \leq C, \quad s \in I_{\varepsilon'}$$

The change of variable $u = s - \pi^2(1-y)/4y$ leads to:

$$\int_0^{+\infty} \frac{\rho_5(s, \varepsilon')}{\sqrt{s + \frac{\pi^2}{4}\varepsilon'}} h\left(\sqrt{s + \frac{\pi^2}{4}\varepsilon'}\right) \leq C \int_{-\frac{3\pi^2}{4}\varepsilon'}^{\frac{\pi^2}{4}\varepsilon'} |\ln u| du. \quad (3.58)$$

Consequently the left hand side in (3.58) goes to 0 as $\varepsilon' \rightarrow 0$. The proof of (2.12) is now complete. \square

3.3. Proof of (2.13) in Theorem 2.1

Our purpose is to get rid off the series that appears in the definition of h . In order to keep the guiding principal, some technical points are proved in Section 3.4. Our proof has five main steps, numbered from 1 to 5.

1) Starting with (2.12), our goal is to simplify

$$\rho(y) := \int_0^{+\infty} \ln \left| 1 - \frac{\pi^2(1-y)}{4ys} \right| \frac{h(\sqrt{s})}{\sqrt{s}} ds.$$

In order to switch the sum and the integral, we need to determine the behavior of h in the vicinity of 0 and $+\infty$.

Lemma 3.13. *One has*

$$(i) \text{ for any } x > 0, |h(x)| \leq 4e^{-2x}/(1 - e^{-2x})^2;$$

$$(ii) \text{ for any } x \in (0, 1), |h(x)| \leq C/\sqrt{x} \text{ where } C \text{ is a positive constant.}$$

The proof of Lemma 3.13 is postponed in Section 3.4. The simplification of ρ is based on the following identity:

$$\ln \left| 1 - \frac{\pi^2(1-y)}{4ys} \right| = \ln \left(\frac{\pi^2(1-y)}{4y} \right) - \ln s + \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right|. \quad (3.59)$$

This gives rise to a decomposition of $\rho(y)$ as the sum of three terms.

Lemma 3.14. *One has*

$$\rho(y) = a_1 \ln \left(\frac{\pi^2(1-y)}{4y} \right) + a_2 + \tau(y) \quad (3.60)$$

$$\text{where } a_1 := \int_0^{+\infty} \frac{h(\sqrt{s})}{\sqrt{s}} ds, \quad a_2 := - \int_0^{+\infty} \ln(s) \frac{h(\sqrt{s})}{\sqrt{s}} ds, \quad (3.61)$$

$$\text{and } \tau(y) := \int_0^{+\infty} \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \frac{h(\sqrt{s})}{\sqrt{s}} ds. \quad (3.62)$$

We calculate τ (Lemma 3.19), a_1 (item **3**) and a_2 (Lemma 3.21).

2) We begin with τ . We can eliminate the sum in $h(\sqrt{s})$ but only on $[\varepsilon, +\infty[$. Let us introduce

$$\varphi(a, n) := \prod_{i=1}^n \left(\frac{a^2}{k^2} - 1 \right)^{(-1)^{k+1}}, \quad a > 0, \quad n \geq 1. \quad (3.63)$$

Lemma 3.15. *For any $y > 0$, $\tau(y) = \lim_{\varepsilon \rightarrow 0} \tau_1(\varepsilon, y)$ where:*

$$\tau_1(\varepsilon, y) := \int_{\varepsilon}^{+\infty} \ln \left| \varphi \left(\sqrt{\frac{4yt}{\pi^2(1-y)}}, \left\lfloor \sqrt{\frac{t}{\varepsilon}} \right\rfloor \right) \right| \frac{dt}{\sqrt{t} \cosh^2(\sqrt{t})}. \quad (3.64)$$

Proof of Lemma 3.15 a) First, using (ii) of Lemma 3.13, we get that the function $s \mapsto \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \frac{h(\sqrt{s})}{\sqrt{s}}$ is integrable on $[0, \infty[$, consequently $\tau(y) = \lim_{\varepsilon \rightarrow 0} \tau_1(\varepsilon, y)$ with

$$\tau_1(\varepsilon, y) = \int_{\varepsilon}^{+\infty} \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \frac{h(\sqrt{s})}{\sqrt{s}} ds. \quad (3.65)$$

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

b) Second, it is possible to permute the integral and the sum in $\tau_1(\varepsilon, y)$ (Equation (3.65)). In that view, we will use the following result: let $(f_k)_{k \geq 1}$ be a sequence of functions defined on $[0, +\infty[$ and such that

$$\sum_{k \geq 1} \int_0^{+\infty} |f_k(s)| ds < +\infty. \quad (3.66)$$

Then

$$\int_0^{+\infty} \left(\sum_{k \geq 1} f_k(s) \right) ds = \int_0^{+\infty} \left(\sum_{k \geq 1} f_k \left(\frac{t}{k^2} \right) \frac{1}{k^2} \right) dt. \quad (3.67)$$

Now let us define

$$f_k(s) := (-1)^{k+1} \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \frac{k}{\sqrt{s} \cosh^2(k\sqrt{s})} \mathbb{1}_{\{s \geq \varepsilon\}}.$$

Since $\cosh x \geq e^x/2$ and $\sqrt{s} \geq \sqrt{\varepsilon}$, we get

$$\begin{aligned} \sum_{k \geq 1} |f_k(s)| &\leq \left| \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \right| \frac{4}{\sqrt{\varepsilon}} \sum_{k \geq 1} k e^{-2k\sqrt{s}} \\ &\leq \frac{4}{\sqrt{\varepsilon}} \left| \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \right| \frac{e^{-2\sqrt{s}}}{(1 - e^{-2\sqrt{\varepsilon}})^2} \end{aligned}$$

using

$$\sum_{k \geq 1} k \rho^k = \rho \frac{d}{d\rho} \left(\sum_{k \geq 1} \rho^k \right) = \frac{\rho}{(1-\rho)^2}.$$

Noticing that $\int_0^{+\infty} e^{-2\sqrt{s}} \left| \ln \left| \frac{4ys}{\pi^2(1-y)} - 1 \right| \right| ds < \infty$, (3.66) holds, we may apply (3.67):

$$\tau_1(\varepsilon, y) = \int_0^{+\infty} \left(\sum_{k \geq 1} (-1)^{k+1} \ln \left| \frac{4yt}{\pi^2(1-y)k^2} - 1 \right| \mathbb{1}_{\{k^2 \leq \frac{t}{\varepsilon}\}} \right) \frac{1}{\sqrt{t} \cosh^2(\sqrt{t})} dt.$$

Since $k \geq 1$, $k^2 \varepsilon \leq t$ implies $t \geq \varepsilon$ and $k^2 \varepsilon \leq t \Leftrightarrow k \leq \sqrt{t}/\varepsilon \Leftrightarrow k \leq \lfloor \sqrt{t}/\varepsilon \rfloor$. We introduce $a := \sqrt{4yt}/(\pi^2(1-y))$, $n := \lfloor \sqrt{t}/\varepsilon \rfloor$, then

$$\sum_{k=1}^n (-1)^{k+1} \ln \left| \frac{a^2}{k^2} - 1 \right| = \ln |\varphi(a, n)|$$

from which we deduce the Lemma 3.15. \square

To determine the limit of $\tau_1(y, \varepsilon)$ in (3.65), we calculate the limit of $\varphi(a, m)$ when $m \rightarrow \infty$ (cf. Lemma 3.17). We prove in Lemma 3.18 that the dominated convergence theorem applies. We begin with modifying $\varphi(a, m)$ so that its limit as $m \rightarrow \infty$ can be calculated.

Lemma 3.16. *One has*

$$\varphi(a, 2n) = \frac{\Gamma(n+1 - (a+1)/2)\Gamma(n+1 + (a-1)/2)}{\Gamma(n+1 - a/2)\Gamma(n+1 + a/2)} \left(\frac{\Gamma(n+1)}{\Gamma(n+1/2)} \right)^2 \alpha(a)$$

where

$$\alpha(a) := \pi \frac{\Gamma(1 - a/2)\Gamma(1 + a/2)}{\Gamma(1 - (a+1)/2)\Gamma(1 + (a-1)/2)}.$$

Proof of Lemma 3.16 By definition, we get

$$\varphi(a, 2n) = \prod_{k=1}^n \left(\frac{k - \frac{a+1}{2}}{k - \frac{a}{2}} \right) \prod_{k=1}^n \left(\frac{k + \frac{a-1}{2}}{k + \frac{a}{2}} \right) \left(\prod_{k=1}^n \frac{k}{k - \frac{1}{2}} \right)^2.$$

Recalling that $\Gamma(x+1) = x\Gamma(x)$ for all $x \notin \mathbb{N}$, we easily deduce

$$\prod_{k=1}^n (k+b) = \frac{\Gamma(n+1+b)}{\Gamma(1+b)}$$

and

$$\varphi(a, 2n) = \frac{\Gamma(n+1 - (1+a)/2)\Gamma(n+1 + (a-1)/2)}{\Gamma(n+1 - a/2)\Gamma(n+1 + a/2)} \left(\frac{\Gamma(n+1)}{\Gamma(n+1/2)} \right)^2 \alpha(a)$$

with $\alpha(a) = (\Gamma(1/2))^2 \Gamma(1 - \frac{a}{2})\Gamma(1 + \frac{a}{2}) / (\Gamma(1 - \frac{1+a}{2})\Gamma(1 + \frac{a-1}{2}))$. It remains to notice that $\Gamma(1/2) = \sqrt{\pi}$ to conclude the proof. \square

We are now able to compute the limit of $\varphi(a, m)$ as m goes to infinity.

Lemma 3.17. *For any a ,*

$$\lim_{m \rightarrow +\infty} |\varphi(a, m)| = |\alpha(a)| = \left| \frac{\pi a}{2} \cot \left(\frac{\pi a}{2} \right) \right|.$$

Proof of Lemma 3.17 (i) Assume that $m = 2n$. Recall that $\Gamma(1+x) \sim \sqrt{2\pi}x^{x+1/2}e^{-x}$ as $x \rightarrow \infty$ and notice that

$$\frac{n - (a+1)/2}{n - a/2} = 1 - \frac{1}{2n} + o(1/n) \quad \text{and} \quad (n-a/2) \ln \left(\frac{n - (a+1)/2}{n - a/2} \right) \rightarrow -1/2.$$

Then

$$\frac{\Gamma(n+1 - (a+1)/2)}{\Gamma(n+1 - a/2)} \sim 1/\sqrt{n}, n \rightarrow \infty.$$

Changing the variable a in $-a$ leads to

$$\frac{\Gamma(n+1 + (a-1)/2)}{\Gamma(n+1 + a/2)} \sim 1/\sqrt{n}, n \rightarrow \infty.$$

Similarly, we get

$$\frac{\Gamma(n+1)}{\Gamma(n+1/2)} \sim \sqrt{n}, n \rightarrow \infty.$$

Consequently, $\lim_{n \rightarrow +\infty} \varphi(a, 2n) = \alpha(a)$

(ii) Assume now that $m = 2n+1$. It suffices to notice that $\varphi(a, 2n+1) = \varphi(a, 2n) \left(\frac{a^2}{(2n+1)^2} - 1 \right)$ and to apply the previous result to get that the limit of $\varphi(a, 2n+1)$ is $-\alpha(a)$ when $n \rightarrow +\infty$.

(iii) Finally, it remains to simplify $\alpha(a)$. Using the identities $\Gamma(1+x) = x\Gamma(x)$ with $x = a/2$ and $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$, $z \notin -\mathbb{N}$ for $z = a/2$ and $z = 1/2 - a/2$ yields the required result. \square

Since $\left\lfloor \sqrt{t/\varepsilon} \right\rfloor \rightarrow +\infty$ as $\varepsilon \rightarrow 0$, then Lemma 3.17 implies:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \ln \left| \varphi \left(\sqrt{\frac{4ty}{\pi^2(1-y)}}, \left\lfloor \sqrt{\frac{t}{\varepsilon}} \right\rfloor \right) \right| &= \ln \left| \alpha \left(\sqrt{\frac{4ty}{\pi^2(1-y)}} \right) \right| \\ &= \ln \left(\sqrt{\frac{ty}{1-y}} \left| \cot \left(\sqrt{\frac{ty}{1-y}} \right) \right| \right). \end{aligned}$$

It remains to apply the dominated convergence theorem to get the limit of $\tau_1(y, \varepsilon)$. Its legal use is justified by the following lemma.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Lemma 3.18. (i) First,

$$\left| \ln \left| \varphi \left(\sqrt{\frac{4yt}{\pi^2(1-y)}}, \left\lfloor \sqrt{\frac{t}{\varepsilon}} \right\rfloor \right) \right| \right| \leq \varphi_1(t)$$

with

$$\varphi_1(t) := \sum_{k \geq 1} \left| \ln \left| 1 - \frac{t}{a_3 k^2} \right| \right| \quad \text{and} \quad a_3 := \frac{\pi^2(1-y)}{4y}.$$

(ii) Second,

$$A := \int_0^\infty \varphi_1(t) \frac{dt}{\sqrt{t} \cosh^2(\sqrt{t})} < \infty. \quad (3.68)$$

Proof of Lemma 3.18

(i) By definition, we have $\ln |\varphi(a, n)| = \sum_{k=1}^n (-1)^{k+1} \ln |1 - a^2/k^2|$ and thus

$$\ln |\varphi(a, n)| \leq \sum_{k=1}^n \left| \ln \left| 1 - \frac{a^2}{k^2} \right| \right|.$$

(ii) We decompose A defined by (3.68) as $A = A_1 + A_2$ with

$$A_1 = \int_0^{a_3/2} \varphi_1(t) \frac{dt}{\sqrt{t} \cosh^2(\sqrt{t})} \quad \text{and} \quad A_2 = \int_{a_3/2}^\infty \varphi_1(t) \frac{dt}{\sqrt{t} \cosh^2(\sqrt{t})}.$$

We prove that A_1 and A_2 are finite.

• Since there exists $c > 0$ such that $|\ln(1-x)| \leq cx$ for $0 < x < 1/2$, condition $t \leq a_3/2$ implies that

$$\left| \ln \left| 1 - \frac{t}{a_3 k^2} \right| \right| \leq \frac{ct}{a_3 k^2}.$$

from which we deduce that

$$A_1 \leq \frac{c}{a_3} \left(\sum_{k \geq 1} \frac{1}{k^2} \right) \int_0^{a_3/2} \left(\frac{\sqrt{t} dt}{\cosh^2(\sqrt{t})} \right) < \infty.$$

• Since $\varphi_1(t)$ is a series with positive terms, we have:

$$A_2 = \sum_{k \geq 1} \int_{a_3/2}^\infty \left| \ln \left| 1 - \frac{t}{a_3 k^2} \right| \right| \frac{dt}{\sqrt{t} \cosh^2(\sqrt{t})}.$$

and the change of variable $t = k^2s$ leads to:

$$A_2 = \sum_{k \geq 1} k \int_{a_3/2k^2}^{\infty} \left| \ln \left| 1 - \frac{s}{a_3} \right| \right| \frac{ds}{\sqrt{s} \cosh^2(k\sqrt{s})}.$$

Now since $\cosh x \geq e^x/2$, we get

$$\begin{aligned} A_2 &\leq 4 \int_0^{\infty} \left| \ln \left| 1 - \frac{s}{a_3} \right| \right| \frac{1}{\sqrt{s}} \left(\sum_{k \geq 1} k e^{-2k\sqrt{s}} \right) ds \\ &\leq 4 \int_0^{\infty} \left| \ln \left| 1 - \frac{s}{a_3} \right| \right| \frac{e^{-2\sqrt{s}}}{(1 - e^{-2\sqrt{s}})^2} \frac{ds}{\sqrt{s}} \end{aligned}$$

Finally we conclude that the integral converges at infinity and 0. \square

We have established the following lemma.

Lemma 3.19. *The function τ can be simplified as:*

$$\tau(y) = 2 \int_0^{+\infty} \ln \left(s \sqrt{\frac{y}{1-y}} \left| \cot \left(s \sqrt{\frac{y}{1-y}} \right) \right| \right) \frac{ds}{\cosh^2(s)}. \quad (3.69)$$

3) Let us prove that $a_1 = 1$. We use the following probabilistic result.

Lemma 3.20. *Let $(\zeta(\alpha))_{0 < \alpha < 1}$ be a sequence of geometrical random variables with parameter α . Then $\eta(\alpha) := (1 - \alpha)^{\zeta(\alpha)}$ converges in distribution to $\mathcal{U}([0, 1])$ as $\alpha \rightarrow 0$.*

Admit for a while Lemma 3.20 whose proof is postponed at the end of item **3)**.

First, by Lemma 3.13,

$$a_1 = \lim_{\varepsilon \rightarrow 0} a_1(\varepsilon) \quad \text{where} \quad a_1(\varepsilon) := \int_{\varepsilon}^{+\infty} \frac{h(\sqrt{s})}{\sqrt{s}} ds.$$

By permuting the sum and the integral, letting $k\sqrt{s} = t$ and introducing

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

$\nu = e^{-2k\sqrt{\varepsilon}}$, we obtain:

$$\begin{aligned} a_1(\varepsilon) &= 2 \sum_{k=1}^{+\infty} (-1)^{k+1} \int_{k\sqrt{\varepsilon}}^{+\infty} \frac{dt}{\cosh^2 t} = 2 \sum_{k=1}^{+\infty} (-1)^{k+1} (1 - \tanh(k\sqrt{\varepsilon})) \\ &= 4 \sum_{k=1}^{+\infty} (-1)^{k+1} \frac{e^{-2k\sqrt{\varepsilon}}}{1 + e^{-2k\sqrt{\varepsilon}}} = 4 \sum_{k=1}^{+\infty} (-1)^{k+1} \frac{\nu^k}{1 + \nu^k} \\ &= 4 \sum_{k=1}^{+\infty} \frac{\nu^{2k-1}(1 - \nu)}{(1 + \nu^{2k-1})(1 + \nu^{2k})}, \end{aligned}$$

where the last equality comes from separating the odd and even numbers. Consequently,

$$a_1(\varepsilon) = \frac{4\nu}{1 + \nu} \sum_{k=1}^{+\infty} \frac{1}{(1 + \frac{1}{\nu}\nu^{2k})(1 + \nu^{2k})} \mathbb{P}(\zeta(1 - \nu^2) = k) = \frac{4\nu}{1 + \nu} \mathbb{E} \left[\varphi \left(\frac{1}{\nu}, \tau(1 - \nu^2) \right) \right]$$

$$\text{where } \varphi(u, v) = \frac{1}{(1 + uv)(1 + v)}, \quad u, v \geq 0.$$

Since φ is continuous, bounded and $\nu \rightarrow 1$ as $\varepsilon \rightarrow 0$, it remains to apply Lemma 3.20 to get

$$a_1 = \lim_{\varepsilon \rightarrow 0} a_1(\varepsilon) = 2 \int_0^1 \frac{du}{(1 + u)^2} = 1. \quad \square$$

Proof of Lemma 3.20 Let $0 < u < 1$. Then

$$\mathbb{P}(\eta(\alpha) < u) = \mathbb{P} \left(\zeta(\alpha) > \frac{\ln u}{\ln(1 - \alpha)} \right) = (1 - \alpha)^{\lfloor \ln u / \ln(1 - \alpha) \rfloor} \rightarrow u, \quad \text{as } \alpha \rightarrow 0. \quad \square$$

4) We calculate a_2 . We admit two intermediate results (whose proofs are postponed in Section 3.4) so that the reader can have quickly the explicit value of a_2 .

Lemma 3.21. *We have*

$$a_2 = -2 \int_0^{+\infty} \frac{\ln t}{\cosh^2(t)} dt - \ln \left(\frac{\pi^2}{4} \right). \quad (3.70)$$

Proof of Lemma 3.21 We proceed as for the calculation of a_1 and we straightforwardly get

$$a_2 = -\lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{+\infty} \frac{h(\sqrt{s})}{\sqrt{s}} \ln(s) ds = -\lim_{\varepsilon \rightarrow 0} (a_{21}(\varepsilon) + a_{22}(\varepsilon))$$

where

$$\begin{aligned} a_{21}(\varepsilon) &:= 4 \sum_{k=1}^{+\infty} \int_{(2k-1)\sqrt{\varepsilon}}^{2k\sqrt{\varepsilon}} \frac{\ln t}{\cosh^2(t)} ds, \\ a_{22}(\varepsilon) &:= \sum_{k=1}^{+\infty} \left(\ln(2k) \int_{2k\sqrt{\varepsilon}}^{+\infty} \frac{dt}{\cosh^2(t)} - \ln(2k-1) \int_{(2k-1)\sqrt{\varepsilon}}^{+\infty} \frac{dt}{\cosh^2(t)} \right). \end{aligned} \quad (3.71)$$

We prove in Section 3.4 the following results

$$\lim_{\varepsilon \rightarrow 0} a_{21}(\varepsilon) = 2 \int_0^{+\infty} \frac{\ln t}{\cosh^2(t)} dt, \quad (3.72)$$

$$\lim_{\varepsilon \rightarrow 0} a_{22}(\varepsilon) = \ln \left(\frac{\pi^2}{4} \right). \quad (3.73)$$

Then, (3.70) follows immediately. \square

5) Finally, we are able to compute $\rho(y)$. Plugging (3.69) into (3.60), using $a_1 = 1$ and (3.70), we get

$$\begin{aligned} \rho(y) &= \ln \left(\frac{1-y}{y} \right) - 2 \int_0^{+\infty} \ln s \frac{ds}{\cosh^2(s)} \\ &\quad + 2 \int_0^{+\infty} \ln \left| s \sqrt{\frac{y}{1-y}} \cot \left(s \sqrt{\frac{y}{1-y}} \right) \right| \frac{ds}{\cosh^2(s)}. \end{aligned}$$

Note that $\int_0^{+\infty} ds / \cosh^2(s) = 1$, consequently:

$$\rho(y) = 2 \int_0^{+\infty} \ln \left| \cot \left(s \sqrt{\frac{y}{1-y}} \right) \right| \frac{ds}{\cosh^2(s)}.$$

Setting $x = s \sqrt{y/(1-y)}$, we have:

$$\rho(y) = 2 \sqrt{\frac{1-y}{y}} \int_0^{+\infty} \ln |\cot x| \frac{ds}{\cosh^2 \left(x \sqrt{\frac{1-y}{y}} \right)}.$$

We conclude to (2.13) and the proof Theorem of 2.1 .

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

3.4. Technical results

In this subsection, we prove the intermediate results used in Section 3, namely Lemma 3.13, and identities (3.72) and (3.73).

3.4.1. Proof of Lemma 3.13

(i) The first inequality is direct as soon as one notices that $\cosh x \geq e^x/2$.
(ii) Let h_0 and k be the functions defined by $h_0(x) := xh(x)$ for $x > 0$ and $h_1(x) := x/\cosh^2(x)$ for $x \geq 0$. The Taylor expansion of h_1 with integral rest can be written in the following form: $h_1(b) - h_1(a) = (b - a)h_1'(b) + R(a, b)$ where $R(a, b) := (b - a) \int_0^1 (h_1'(a + r(b - a)) - h_1'(b))dr$. Then,

$$h_0(x) = \sum_{k \geq 1} (-1)^{k+1} h_1(kx) = -(h_{01}(x) + h_{02}(x) + h_{03}(x))$$

where

$$\begin{aligned} h_{01}(x) &= x \left(\sum_{1 \leq n \leq x^*} h_1'(2nx) \right) \\ h_{02}(x) &= \sum_{1 \leq n \leq x^*} R((2n - 1)x, 2nx) \\ h_{03}(x) &= \sum_{n > x^*} (h_1(2nx) - h_1((2n - 1)x)) \end{aligned}$$

and $x^* = x^{-3/2}$. The derivative of h_1 being bounded, we deduce that $|h_{02}(x)| \leq c\sqrt{x}$. Since $h_1(x) \leq ce^{-x}$, we can easily prove that $|h_{03}(x)| \leq ce^{-2/\sqrt{x}}/x$. As for $h_{01}(x)$, we decompose this sum with the related integral:

$$h_{01}(x) = \int_0^{\bar{x}} h_1'(2y)dy + \tilde{h}_{01}(x)$$

where $\tilde{h}_{01}(x) = x \left(\sum_{1 \leq n \leq n_x} \int_{(n-1)x}^{nx} (h_1'(2nx) - h_1'(2y))dy \right)$, $n_x = \lfloor x^* \rfloor$ and $\bar{x} = n_x x$. From the previous estimates, we have $|\tilde{h}_{01}(x)| \leq c\sqrt{x}$ and

$$\left| \int_0^{\bar{x}} h_1'(2y)dy \right| = \left| \frac{1}{2} h_1(\bar{x}) \right| \leq ce^{-2/\sqrt{\bar{x}}}.$$

□

3.4.2. Proof of Equation (3.72)

We define $f(t) := \ln t / \cosh^2(t)$ and

$$a'_{21}(\varepsilon) := 4 \sum_{k=1}^{+\infty} \int_{2k\sqrt{\varepsilon}}^{(2k+1)\sqrt{\varepsilon}} f(t) dt.$$

- Obviously, we first have

$$a_{21}(\varepsilon) + a'_{21}(\varepsilon) = 4 \int_{\sqrt{\varepsilon}}^{+\infty} f(t) dt$$

$$a_{21}(\varepsilon) - a'_{21}(\varepsilon) = 4 \sum_{k=1}^{+\infty} \int_{(2k-1)\sqrt{\varepsilon}}^{2k\sqrt{\varepsilon}} (f(t) - f(t + \sqrt{\varepsilon})) dt.$$

- Second, we prove that $\lim_{\varepsilon \rightarrow 0} (a_{21}(\varepsilon) - a'_{21}(\varepsilon)) = 0$. Indeed, we notice that for small values of ε , we have $f(t + \sqrt{\varepsilon}) - f(t) \geq 0$, $t \in [0, 1]$ since for all $t \in [0, 1]$, $f'(t) > 0$. Then

$$|a_{21}(\varepsilon) - a'_{21}(\varepsilon)| \leq 4 \int_0^1 (f(t + \sqrt{\varepsilon}) - f(t)) dt + 4 \int_1^{+\infty} |f(t + \sqrt{\varepsilon}) - f(t)| dt.$$

On one hand,

$$\int_0^1 (f(t + \sqrt{\varepsilon}) - f(t)) dt = \int_1^{1+\sqrt{\varepsilon}} f(t) dt - \int_0^{\sqrt{\varepsilon}} f(t) dt \xrightarrow{\varepsilon \rightarrow 0} 0$$

since f is integrable on \mathbb{R}_+ .

On the other hand, for any $t \geq 1$, $|f'(t)| \leq C e^{-t}$ that leads to

$$\int_1^{+\infty} |f(t + \sqrt{\varepsilon}) - f(t)| dt \leq C \sqrt{\varepsilon} \int_1^{+\infty} e^{-t} dt \xrightarrow{\varepsilon \rightarrow 0} 0.$$

- Finally, we derive (3.72) from the previous results and the following equation

$$a_{21}(\varepsilon) = \frac{1}{2}(a_{21}(\varepsilon) + a'_{21}(\varepsilon)) + \frac{1}{2}(a_{21}(\varepsilon) - a'_{21}(\varepsilon)). \quad \square$$

3.4.3. Proof of Equation (3.73)

It is convenient to introduce $\theta(t) := 1/\cosh^2(t)$, $t \geq 0$ and

$$n = n_{\varepsilon,t} := \left\lfloor \frac{1}{2} \left(1 + \frac{t}{\sqrt{\varepsilon}} \right) \right\rfloor. \quad (3.74)$$

We begin with a new decomposition of $a_{22}(\varepsilon)$.

Lemma 3.22.

$$a_{22}(\varepsilon) = 4 \int_{\sqrt{\varepsilon}}^{+\infty} \left(g_{\varepsilon}^1(t) + g_{\varepsilon}^2(t) + O\left(\frac{1}{n}\right) (\theta(t + \sqrt{\varepsilon}) + \theta(t)) \right) dt \quad (3.75)$$

where

$$g_{\varepsilon}^1(t) = n \ln n (\theta(t + \sqrt{\varepsilon}) - \theta(t)) + \frac{1}{2} \ln n \theta(t + \sqrt{\varepsilon}),$$

$$g_{\varepsilon}^2(t) = (\ln 2 - 1)n (\theta(t + \sqrt{\varepsilon}) - \theta(t)) \ln \sqrt{2\pi} \theta(t + \sqrt{\varepsilon}) - \frac{1}{2} \ln 2 \theta(t).$$

Proof of Lemma 3.22

a) We introduce

$$S_m^e = \sum_{k=1}^m \ln(2k) \quad \text{and} \quad S_m^o = \sum_{k=1}^m \ln(2k-1).$$

Using Stirling formula

$$k! = \sqrt{2\pi} k^{k+1/2} \exp \left\{ -k + \frac{\eta(k)}{12k} \right\}, \quad k \geq 1$$

where $1 - 1/(12k+1) \leq \eta(k) < 1$, we easily deduce the asymptotic behavior of S_m^e and S_m^o as $m \rightarrow +\infty$:

$$S_k^e = k \ln k + (\ln 2 - 1)k + \frac{1}{2} \ln k + \ln \sqrt{2\pi} + O\left(\frac{1}{k}\right), \quad (3.76)$$

$$S_k^o = k \ln k + (\ln 2 - 1)k + \frac{1}{2} \ln 2 + O\left(\frac{1}{k}\right) \quad (3.77)$$

where $O(1/k)$ stands for a sequence bounded by $1/k$.

$$\begin{aligned}
\sum_{k=1}^n \ln(2k) \int_{2k\sqrt{\varepsilon}}^{+\infty} \theta(t) dt &= \sum_{k=1}^n \ln(2k) \int_{(2k-1)\sqrt{\varepsilon}}^{+\infty} \theta(t + \sqrt{\varepsilon}) dt \\
&= \int_{\sqrt{\varepsilon}}^{+\infty} \left(\sum_{k=1}^n \ln(2k) \mathbb{1}_{\{(2k-1)\sqrt{\varepsilon} \leq t\}} \right) \theta(t + \sqrt{\varepsilon}) dt \\
&= \int_{\sqrt{\varepsilon}}^{+\infty} S_n^e \theta(t + \sqrt{\varepsilon}) dt
\end{aligned}$$

b) One has

and similarly,

$$\sum_{k=1}^n \ln(2k-1) \int_{(2k-1)\sqrt{\varepsilon}}^{+\infty} \theta(t) dt = \int_{\sqrt{\varepsilon}}^{+\infty} S_n^o \theta(t) dt.$$

Then (3.75) is a direct consequence of (3.71), (3.76) and (3.77). \square

By Lemma 3.22, $a_{22}(\varepsilon)$ is the sum of three terms. We determine the limit of each term separately.

1) It is clear that

$$\lim_{\varepsilon \rightarrow 0} \int_{\sqrt{\varepsilon}}^{+\infty} O\left(\frac{1}{n}\right) (\theta(t + \sqrt{\varepsilon}) + \theta(t)) dt = 0. \quad (3.78)$$

2) Note that

$$n|\theta(t + \sqrt{\varepsilon}) - \theta(t)| \leq C e^{-t}, \quad t \geq 0, \quad \varepsilon > 0. \quad (3.79)$$

Moreover, $n \sim \frac{t}{2\sqrt{\varepsilon}}$ as ε goes to 0 and thus $n(\theta(t + \sqrt{\varepsilon}) - \theta(t)) \sim \frac{t}{2}\theta'(t)$, $\varepsilon \rightarrow 0$. We apply the Lebesgue's dominated convergence theorem and we get

$$\begin{aligned}
\lim_{\varepsilon \rightarrow 0} \int_{\sqrt{\varepsilon}}^{+\infty} n(\theta(t + \sqrt{\varepsilon}) - \theta(t)) dt &= \frac{1}{2} \int_0^{+\infty} t\theta'(t) dt \\
&= \frac{1}{2} \left([t\theta(t)]_0^{+\infty} - \int_0^{+\infty} \theta(t) dt \right) = -\frac{1}{2}.
\end{aligned}$$

Finally,

$$\lim_{\varepsilon \rightarrow 0} \int_{\sqrt{\varepsilon}}^{+\infty} g_{\varepsilon}^2(t) dt = \frac{1}{2} (\ln \pi + 1 - \ln 2). \quad (3.80)$$

3) We are interested in the limit of the integral involving $g_{\varepsilon}^1(t)$. In that view, we decompose $g_{\varepsilon}^1(t)$ in Lemma 3.23 using a new class of functions that we define as follows. The sequence of functions $(R_{\varepsilon}(t), t \geq 0)_{\varepsilon \in (0, 1/2]}$ is said to belong to the class \mathcal{R} if

$$(i) \lim_{\varepsilon \rightarrow 0} R_{\varepsilon}(t) = 0, \quad t \geq 0, \quad (3.81)$$

$$(ii) \int_0^{+\infty} \sup_{0 < \varepsilon < 1/2} |R_{\varepsilon}(t)| dt < +\infty. \quad (3.82)$$

Clearly, if $(R_{\varepsilon}(t), t \geq 0)_{\varepsilon \in (0, 1/2]}$ belongs to the class \mathcal{R} then $\lim_{\varepsilon \rightarrow 0} \int_0^{+\infty} R_{\varepsilon}(t) dt = 0$.

Lemma 3.23. *One has*

$$g_{\varepsilon}^1(t) = \frac{1}{2} (t \ln(t) \theta'(t) + \ln(t) \theta(t)) - \frac{1}{2} \ln(2\sqrt{\varepsilon}) \left(t \frac{\theta(t + \sqrt{\varepsilon}) - \theta(t)}{\sqrt{\varepsilon}} + \theta(t) \right) + R_{\varepsilon}(t). \quad (3.83)$$

Proof of Lemma 3.23 The function g_{ε}^1 (defined in Lemma 3.22) is the sum of two terms. We study each of them separately.

a) We prove that

$$n \ln n (\theta(t + \sqrt{\varepsilon}) - \theta(t)) = \frac{t}{2\sqrt{\varepsilon}} \ln \left(\frac{t}{2\sqrt{\varepsilon}} \right) (\theta(t + \sqrt{\varepsilon}) - \theta(t)) + R_{\varepsilon}(t) \quad (3.84)$$

where n has been defined in (3.74) and $(R_{\varepsilon}(t), t \geq 0)_{\varepsilon \in (0, 1/2]} \in \mathcal{R}$. It suffices to show that

$$\begin{aligned} R_{\varepsilon}(t) &= n \ln n (\theta(t + \sqrt{\varepsilon}) - \theta(t)) - \frac{t}{2\sqrt{\varepsilon}} \ln \left(\frac{t}{2\sqrt{\varepsilon}} \right) (\theta(t + \sqrt{\varepsilon}) - \theta(t)) \\ &= (\lambda(n) - \lambda(t/2\sqrt{\varepsilon})) (\theta(t + \sqrt{\varepsilon}) - \theta(t)) \end{aligned} \quad (3.85)$$

belongs to \mathcal{R} where $\lambda(x) = x \ln x$.

Using the definition (3.74) of n , we have $\left|n - \frac{t}{2\sqrt{\varepsilon}}\right| \leq \frac{1}{2}$. Then

$$\left|\lambda(n) - \lambda\left(\frac{t}{2\sqrt{\varepsilon}}\right)\right| \leq \frac{1}{2} \sup_{u \in J} |\lambda'(u)| \leq \frac{1}{2} \left(1 + \ln\left(1 + \frac{t}{\sqrt{\varepsilon}}\right)\right) \quad (3.86)$$

where J is the interval with end points n and $t/2\sqrt{\varepsilon}$. It is clear that

$$|\theta(t + \sqrt{\varepsilon}) - \theta(t)| \leq C\sqrt{\varepsilon}e^{-t}. \quad (3.87)$$

Then (3.81) and (3.82) follow from (3.85), (3.86) and (3.87).

b) Since $|\ln a - \ln b| \leq |a - b|/\min\{a, b\}$, $a, b > 0$, we have $\lim_{\varepsilon \rightarrow 0} (\ln n - \ln(t/2\sqrt{\varepsilon})) = 0$ using (3.74). Using moreover (3.87), we deduce that

$$\ln(n)\theta(t + \sqrt{\varepsilon}) = \ln\left(\frac{t}{2\sqrt{\varepsilon}}\right)\theta(t) + R_\varepsilon(t). \quad (3.88)$$

c) Then (3.83) follows from (3.84), (3.88), the following identity

$$\frac{t \ln t}{2\sqrt{\varepsilon}} (\theta(t + \sqrt{\varepsilon}) - \theta(t)) = \frac{t \ln t}{2} \theta'(t) + R_\varepsilon(t)$$

and easy calculations. □

According to Lemma 3.23,

$$\int_{\sqrt{\varepsilon}}^{+\infty} g_\varepsilon^1(t) dt = \theta_1 + \theta_2(\varepsilon) + o(\varepsilon) \quad (3.89)$$

where

$$\theta_1 := \frac{1}{2} \int_0^{+\infty} (t \ln(t)\theta'(t) + \ln(t)\theta(t)) dt, \quad (3.90)$$

$$\theta_2(\varepsilon) := -\frac{1}{2} \ln(2\sqrt{\varepsilon}) \int_{\sqrt{\varepsilon}}^{+\infty} \left(t \frac{\theta(t + \sqrt{\varepsilon}) - \theta(t)}{\sqrt{\varepsilon}} + \theta(t)\right) dt. \quad (3.91)$$

We calculate θ_1 and the limit of $\theta_2(\varepsilon)$ as $\varepsilon \rightarrow 0$.

(i) By an integration by parts, it comes

$$\int_0^{+\infty} t \ln(t)\theta'(t) dt = - \int_0^{+\infty} \theta(t) dt - \int_0^{+\infty} \ln(t)\theta(t) dt = -1 - \int_0^{+\infty} \ln(t)\theta(t) dt$$

1
2
3
4
5
6
7
8
9 that implies $\theta_1 = -1/2$.

10
11 **(ii)** We determine the limit of $\theta_2(\varepsilon)$, $\varepsilon \rightarrow 0$.

12 Noticing that

13
14
15
$$\int_{\sqrt{\varepsilon}}^{+\infty} t \frac{\theta(t + \sqrt{\varepsilon}) - \theta(t)}{\sqrt{\varepsilon}} dt = \frac{1}{\sqrt{\varepsilon}} \left(\int_{2\sqrt{\varepsilon}}^{+\infty} (t - \sqrt{\varepsilon})\theta(t) dt - \int_{\sqrt{\varepsilon}}^{+\infty} t\theta(t) dt \right)$$

16
17
$$= - \int_{\sqrt{\varepsilon}}^{+\infty} \theta(t) dt - \frac{1}{\sqrt{\varepsilon}} \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} (t - \sqrt{\varepsilon})\theta(t) dt.$$

18
19
20
21

22 Consequently,

23
24
$$\theta_2(\varepsilon) = \frac{1}{2\sqrt{\varepsilon}} \ln(2\sqrt{\varepsilon}) \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} (t - \sqrt{\varepsilon})\theta(t) dt$$

25
26
27

28 and $\lim_{\varepsilon \rightarrow 0} \theta_2(\varepsilon) = 0$. Moreover using (3.89) and $\theta_1 = -1/2$, one gets

29
30
$$\lim_{\varepsilon \rightarrow 0} \int_{\sqrt{\varepsilon}}^{+\infty} g_{\varepsilon}^1(t) dt = -\frac{1}{2}. \quad (3.92)$$

31
32
33

34 Finally, plugging the limits (3.78), (3.80) and (3.92) into (3.75) implies (3.73)
35 which completes the proof. \square
36
37

38 References

- 39
40 [1] P. Billingsley, Convergence of probability measures, Wiley Series in
41 Probability and Statistics: Probability and Statistics, Second Edition,
42 A Wiley-Interscience Publication, John Wiley & Sons Inc., New York,
43 1999.
44
45 [2] P. Biane, J. Pitman, M. Yor, Probability laws related to the Jacobi theta
46 and Riemann zeta functions, and Brownian excursions, Bull. Amer.
47 Math. Soc. (N.S.), American Mathematical Society. Bulletin. New Se-
48 ries, (2001) 38:4 ,435–465 (electronic).
49
50 [3] A.N. Borodin and P. Salminen, Handbook of Brownian motion—
51 facts and formulae, Probability and its Applications, Second Edition,
52 Birkhäuser Verlag, Basel, 2002.
53
54
55
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4
5
6
7
8
9
10 [4] C. Chabriac, A. Lagnoux, S. Mercier, P. Vallois, Elements related to the
11 largest complete excursion of a reflected Brownian motion stopped at
12 a fixed time. Application to local score. Stochastic Processes and their
13 Applications, (2014) 124(12).
14
- 15 [5] J.J. Daudin, M.P. Etienne, P. Vallois, Asymptotic behavior of the lo-
16 cal score of independent and identically distributed random sequences.
17 Stochastic Process. Appl., (2003) 107(1):1–28.
18
- 19 [6] M.I. Fariello, S. Boitard, S. Mercier, D. Robelin, T. Faraut, C. Arnould,
20 E. Le Bihan-Duval, J. Recoquillay, G. Salin, P. Dehais, F. Pitel, C.
21 Leterrier, M. Sancristobal, Accounting for Linkage Disequilibrium in
22 genome scans for selection without individual genotypes : the local score
23 approach. Molecular Ecology, Volume 26, Issue 14, (2017) 3700–3714.
24
- 25 [7] M. Guedj, D. Robelin, M. Hoebeke, M. Lamarine, J. Wojcik, G. Nuel,
26 Detecting Local High-Scoring Segments: a First-Stage Approach for
27 Genome-Wide Association Studies, SAGEMB, (2006) 5:1.
28
- 29 [8] C. Hassenforder, S. Mercier, Exact Distribution of the Local Score for
30 Markovian Sequences. AISM, (2007) 59(4):741–755.
31
- 32 [9] S. Karlin, S.F. Altschul, Methods for assessing the statistical significance
33 of molecular sequence features by using general scoring schemes. PNAS,
34 (1990) 87:2264–2268.
35
- 36 [10] S. Karlin, A. Dembo, Limit distributions of maximal segmental score
37 among Markov-dependent partial sums, AdAP, (1992) 24:113–140.
38
- 39 [11] A. Lagnoux, S. Mercier, P. Vallois, Probability that the maximum of
40 the reflected Brownian motion over a finite interval $[0, t]$ is achieved by
41 its last zero before t . Electronic Communications in Probability, (2015)
42 20(62):1-9.
43
- 44 [12] A. Lagnoux, S. Mercier, P. Vallois, Statistical significance based on
45 length and position of the local score in a model of i.i.d. sequences,
46 Bioinformatics, (2017) 33:5, 654–660.
47
- 48 [13] S. Mercier, J.J. Daudin, Exact distribution for the local score of one
49 i.i.d. random sequence. Jour. Comp. Biol., (2001) 8(4):373–380.
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

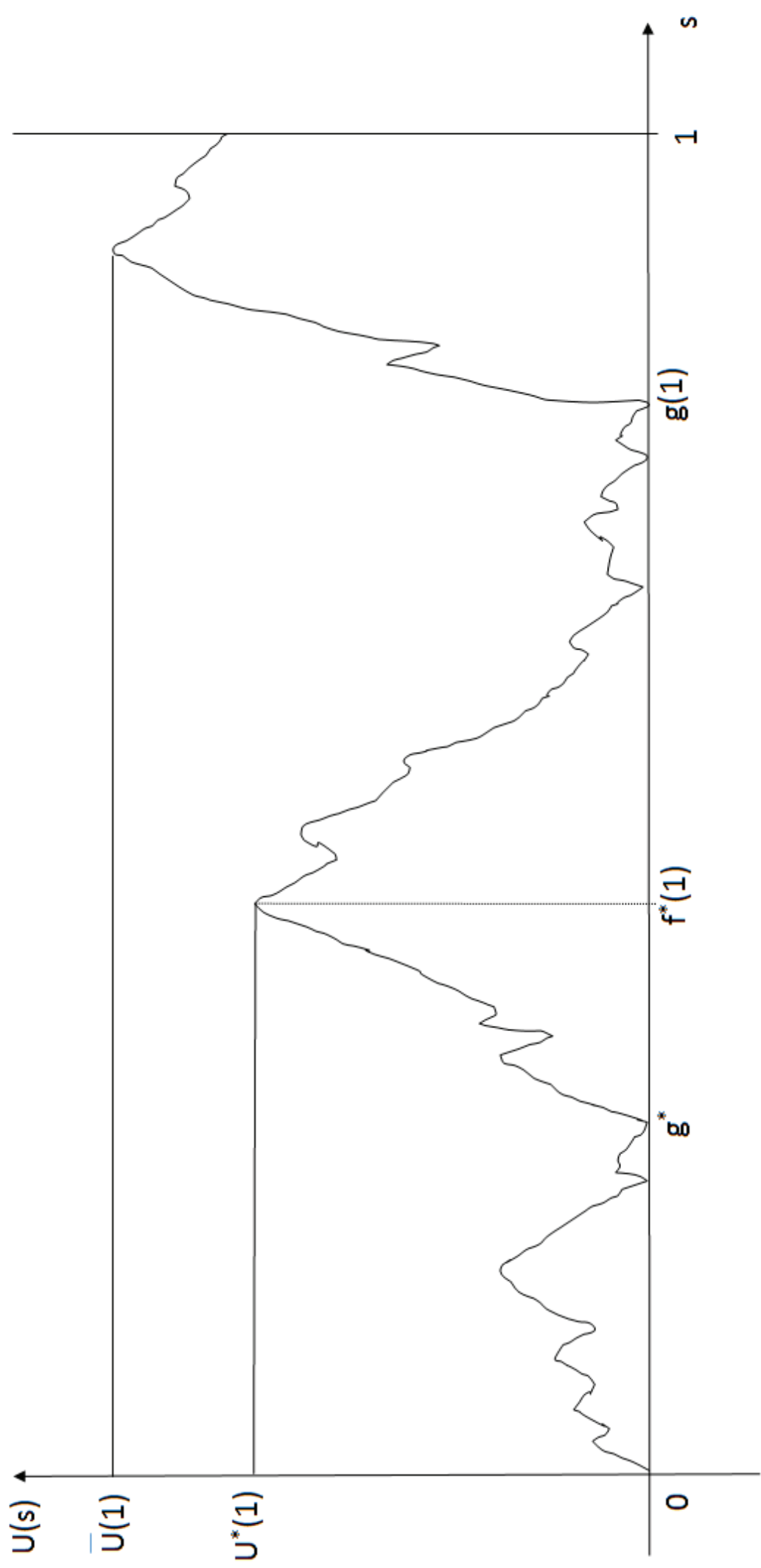
[14] S. Mercier, G. Nuel, A Hidden Markov Model for the Local Score of One Sequence. Submitted.

[15] J. Pitman, M. Yor, Path decompositions of a Brownian bridge related to the ratio of its maximum and amplitude, *Studia Sci. Math. Hungar.*, *Studia Scientiarum Mathematicarum Hungarica. A Quarterly of the Hungarian Academy of Sciences*, (1999) 35:3-4, 457–474.

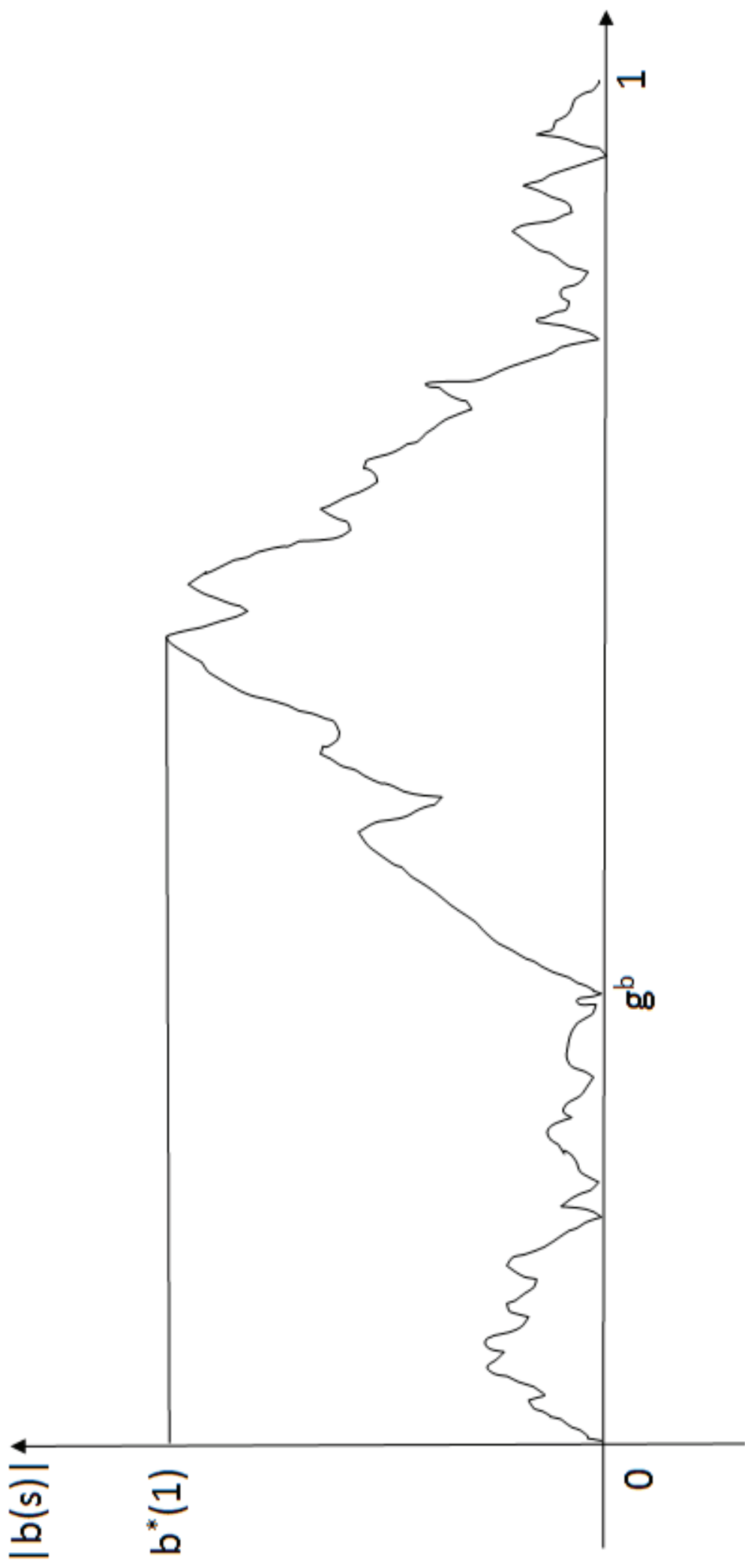
[16] P. Vallois, Sur la loi conjointe du maximum et de l'inverse du temps local du mouvement brownien: application à un théorème de Knight, *Stochastics Stochastics Rep.*, *Stochastics and Stochastics Reports*, (1991) 35:3,175–186.

[17] P. Vallois, Une extension des théorèmes de Ray et Knight sur les temps locaux Browniens, *Probab. Th. Rel. Fields*, (1991) 39:445–482.

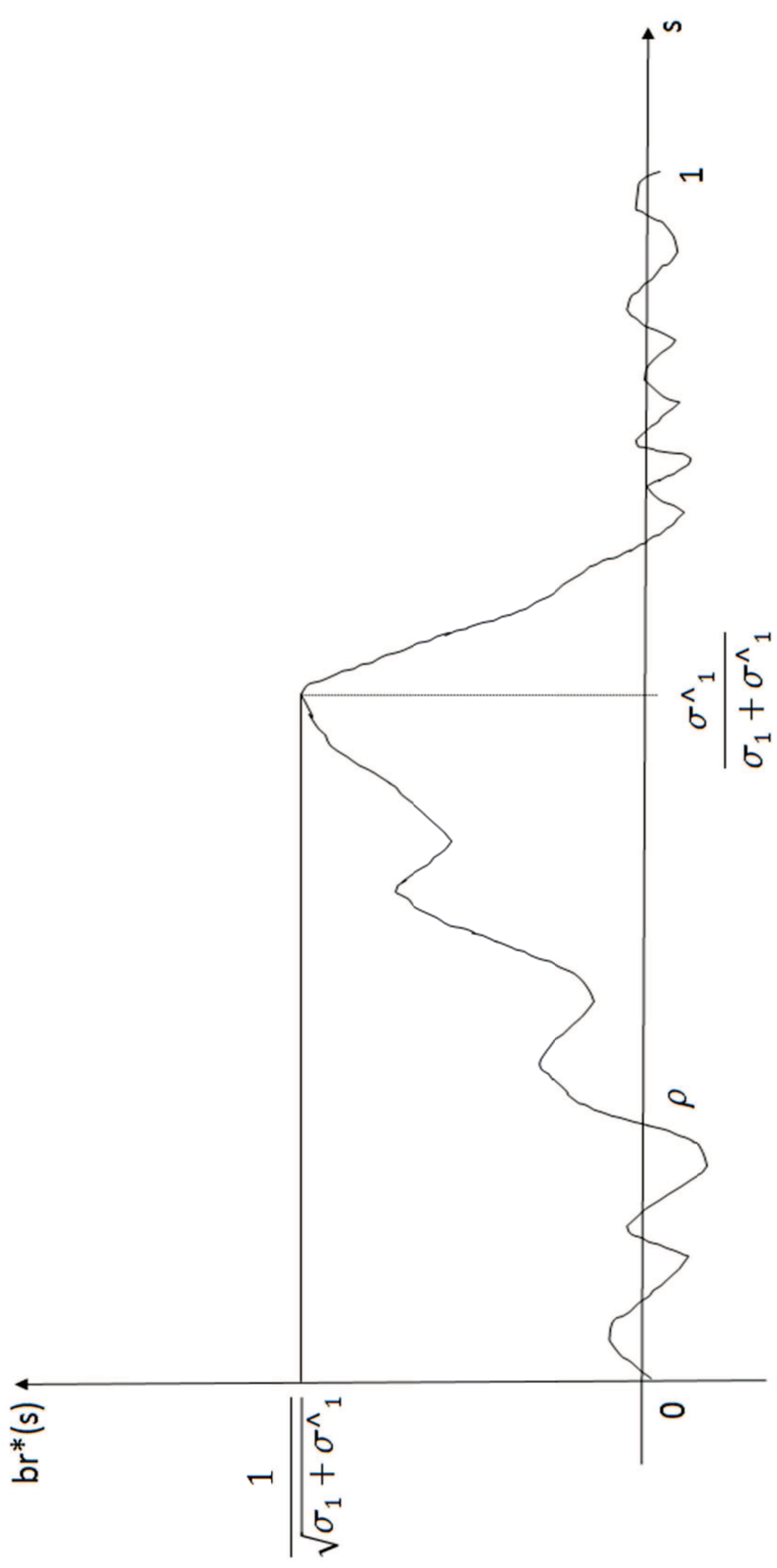
Figure



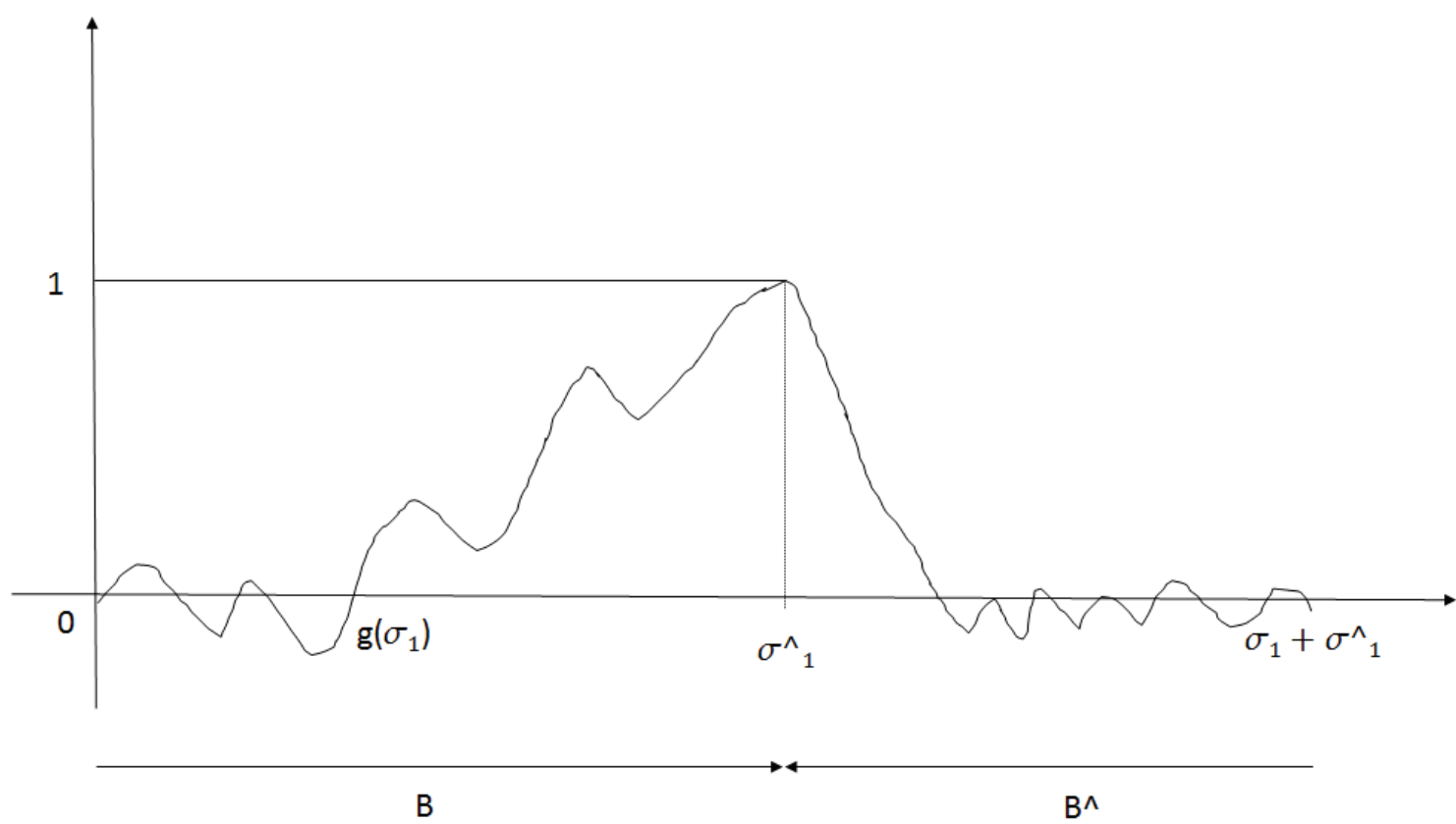
Figure



Figure



Figure



LaTeX Souce Files

[Click here to download LaTeX Souce Files: SPA_Lagnoux_et_al_15jan18.tex](#)