



# Localization of the continuous Anderson Hamiltonian in 1-D

Laure Dumaz<sup>1</sup> · Cyril Labbé<sup>1</sup>

Received: 15 December 2017 / Revised: 25 April 2019 / Published online: 16 May 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

We study the bottom of the spectrum of the Anderson Hamiltonian  $\mathcal{H}_L := -\partial_x^2 + \xi$  on  $[0, L]$  driven by a white noise  $\xi$  and endowed with either Dirichlet or Neumann boundary conditions. We show that, as  $L \rightarrow \infty$ , the point process of the (appropriately shifted and rescaled) eigenvalues converges to a Poisson point process on  $\mathbf{R}$  with intensity  $e^{-x} dx$ , and that the (appropriately rescaled) eigenfunctions converge to Dirac masses located at independent and uniformly distributed points. Furthermore, we show that the shape of each eigenfunction, recentered around its maximum and properly rescaled, is given by the inverse of a hyperbolic cosine. We also show that the eigenfunctions decay exponentially from their localization centers at an explicit rate, and we obtain very precise information on the zeros and local maxima of these eigenfunctions. Finally, we show that the eigenvalues/eigenfunctions in the Dirichlet and Neumann cases are very close to each other and converge to the same limits.

**Keywords** Anderson Hamiltonian · Hill's operator · Localization · Riccati transform · Diffusion

**Mathematics Subject Classification** Primary 60H25 · 60J60; Secondary 35P20

## Contents

1 Introduction	354
2 Convergence of the eigenvalues towards the Poisson point process	366
3 Localization of the eigenfunctions	373
4 Proofs of some first estimates on the diffusion $X_a$	392
5 Crossing of the well	400
6 Neumann boundary conditions	417
References	418

---

✉ Cyril Labbé  
labbe@ceremade.dauphine.fr  
Laure Dumaz  
dumaz@ceremade.dauphine.fr

<sup>1</sup> UMR 7534, CNRS, CEREMADE, Université Paris-Dauphine, PSL Research University, 75016 Paris, France

## 1 Introduction

Consider the Anderson Hamiltonian

$$\mathcal{H}_L = -\partial_x^2 + \xi, \quad x \in (0, L), \quad (1)$$

endowed with either Dirichlet or Neumann boundary conditions. Here, the potential  $\xi$  is taken to be a real white noise on  $(0, L)$ , that is, a mean zero, delta-correlated Gaussian field on  $(0, L)$ . The operator  $\mathcal{H}_L$  is a random Schrödinger operator, sometimes called Hill's operator, that models disordered solids in physics.

There is a competition between the two terms appearing in the operator: while the eigenfunctions of the Laplacian are spread out over the whole box, the multiplication-by- $\xi$  operator tends to concentrate the mass of the eigenfunctions in very small regions. In his seminal article [5], Anderson showed that for a discrete version of the present hamiltonian and in dimension 3, the bottom of the spectrum consists of *localized* eigenfunctions. This phenomenon, now referred to as *Anderson localization*, has been the object of numerous studies, see for instance [22] for an extended survey.

In the present paper, we establish a localization phenomenon at the bottom of the spectrum of  $\mathcal{H}_L$  when  $L \rightarrow \infty$ .

This operator was first studied by the physicist Halperin [19], see also the work of Frisch and Lloyd [13]. The focus in these works was on the macroscopic picture of the eigenvalues in the large  $L$  limit. If  $N(\lambda)$  denotes the number of eigenvalues smaller than  $\lambda$ , the density of states is defined as the derivative of the large  $L$  limit of  $N(\lambda)$  divided by  $L$ . They found that the density of states of the operator  $\mathcal{H}_L$  admits an explicit integral formula, see Equation (11) or Fig. 1.

Later on, Fukushima and Nakao [14] gave a precise formulation of the eigenvalue problem  $\mathcal{H}_L \varphi = \lambda \varphi$ . They proved that for any fixed  $L$ , almost surely  $\mathcal{H}_L$  is a self-adjoint operator on  $L^2(0, L)$ , bounded below, that admits a pure point spectrum

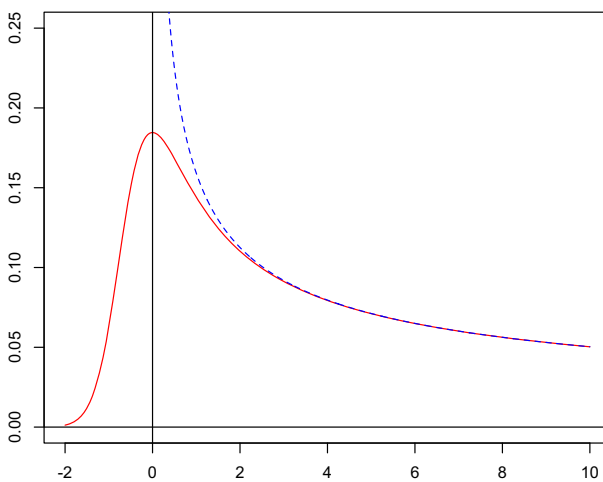


Fig. 1 Density of states of the operator  $\mathcal{H}_L$  (plain line) and of  $u \mapsto -\partial_x^2 u$  (dashed line)

$(\lambda_k)_{k \geq 1}$ , with  $\lambda_1 < \lambda_2 < \dots$ , and that the associated eigenfunctions  $(\varphi_k)_{k \geq 1}$  form an orthonormal basis of  $L^2(0, L)$  and are Hölder  $3/2^-$  (that is, Hölder  $3/2 - \varepsilon$  for all  $\varepsilon > 0$ ). They also rigorously derived the density of states.

Subsequently, McKean [28] established that the first eigenvalue, appropriately shifted and rescaled:

$$-4\sqrt{a_L}(\lambda_1 + a_L), \tag{2}$$

converges as  $L \rightarrow \infty$  to a Gumbel distribution and therefore falls in one of the three famous extreme-value distributions (see also [36] for the asymptotics of the  $k$ -th eigenvalue). The precise definition of  $a_L$  is given right above Equation (12), let us simply mention that as  $L \rightarrow \infty$  we have

$$a_L \sim \left(\frac{3}{8} \ln L\right)^{2/3}.$$

Let us also cite the works [10,11] where the authors obtained an exact formula for the density distribution of the first eigenvalue in terms of an integral over the circular Brownian motion, and a precise asymptotic for its left tail. In these works,  $L$  is fixed and the operator is endowed with periodic boundary conditions.

In all the aforementioned studies, the starting point is the Riccati transform which maps the second order linear differential equation  $\mathcal{H}_L u = \lambda u$  into a first order non-linear one, see below. We will also use this tool in the present paper.

**Statement of our results**

We are interested in the large  $L$  limit behavior of the smallest eigenvalues of  $\mathcal{H}_L$  and of their associated eigenfunctions. We will consider the operator  $\mathcal{H}_L$  endowed with either Dirichlet boundary conditions:  $\varphi(0) = \varphi(L) = 0$ , or Neumann boundary conditions:  $\varphi'(0) = \varphi'(L) = 0$ . In the sequel, if no mention is made of the boundary conditions, then they are taken to be Dirichlet.

Let us introduce the rescaled eigenfunctions:

$$m_k(dt) = L\varphi_k^2(tL)dt, \quad t \in (0, 1), \tag{3}$$

that belong to the space  $\mathcal{P}$  of probability measures on  $[0, 1]$  endowed with the topology of weak convergence. We also introduce  $U_k \in [0, L]$  be the (first, if many) point at which the eigenfunction  $|\varphi_k|$  achieves its maximum.

Our first result shows that asymptotically in  $L$ , the eigenvalues form a Poisson point process on  $\mathbf{R}$  while the eigenfunctions converge to Dirac masses whose locations are uniformly distributed over  $[0, L]$  and independent from the eigenvalues.

**Theorem 1** *The sequence of random variables  $(4\sqrt{a_L}(\lambda_k + a_L), U_k/L, m_k)_{k \geq 1}$  converges in law to  $(\lambda_k^\infty, U_k^\infty, \delta_{U_k^\infty})_{k \geq 1}$  where  $(\lambda_k^\infty, U_k^\infty)_{k \geq 1}$  is a Poisson point process on  $\mathbf{R} \times [0, 1]$  with intensity  $e^x dx \otimes dt$  (where the convergence holds for the set of sequences of elements in  $\mathbf{R} \times [0, 1] \times \mathcal{P}$  endowed with the product topology).*

In the following theorem, we obtain a very precise description of the asymptotic behavior of the eigenfunctions. At a macroscopic level, we show that they decay exponentially fast from their localization centers at rate  $\sqrt{a_L}$ . At a microscopic level, we show that the asymptotic shape around the maximum is given by the inverse of a hyperbolic cosine at a space-scale of order  $(\ln L)^{-1/3}$ .

Let us set for  $i \geq 1$ ,

$$h_i(t) := \frac{\sqrt{2}}{a_L^{1/4}} \left| \varphi_i \left( U_i + \frac{t}{\sqrt{a_L}} \right) \right|, \quad t \in \mathbf{R},$$

where we implicitly set the value of this function to 0 whenever the argument of the function on the right does not belong to  $[0, L]$ . Similarly, we define

$$b_i(t) := \frac{1}{\sqrt{a_L}} \left( B \left( U_i + \frac{t}{\sqrt{a_L}} \right) - B(U_i) \right), \quad t \in \mathbf{R},$$

where  $B(t) := \langle \xi, \mathbf{1}_{[0,t]} \rangle$ ,  $t \in [0, L]$ . Note that  $B$  is a Brownian motion. We also define

$$h(t) := \frac{1}{\cosh(t)}, \quad b(t) := -2 \tanh(t), \quad t \in \mathbf{R}.$$

Finally, denote by  $z_0 := 0 < z_1 < \dots < z_{i-1} < z_i = L$  the zeros of the eigenfunction  $\varphi_i$ .

**Theorem 2** (Exponential decay and shape of the eigenfunctions) *For every  $i \geq 1$ ,*

- **Exponential decay** *There exist deterministic constants  $C_2 > C_1 > 0$  such that with probability going to 1, for all  $t \in [0, L]$ :*

$$C_1 \exp(-(\sqrt{a_L} + \kappa_L)|t - U_i|) \mathbf{1}_{\{t \in D\}} \leq \left| \frac{\varphi_i(t)}{\varphi_i(U_i)} \right| \leq C_2 \exp(-(\sqrt{a_L} - \kappa_L)|t - U_i|), \tag{4}$$

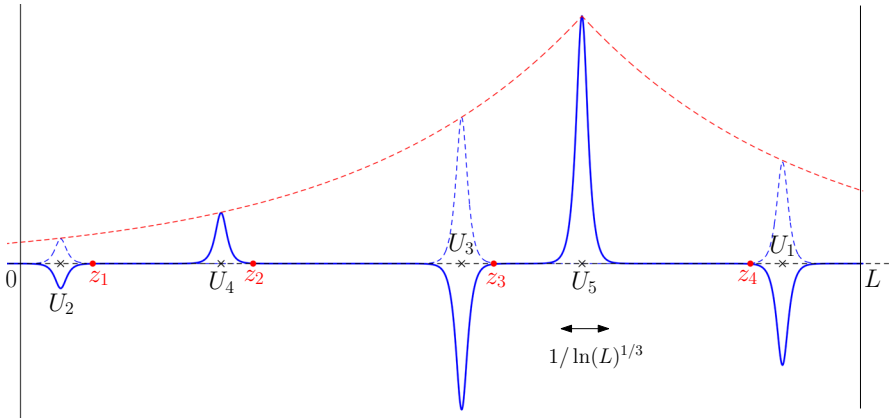
where

$$\kappa_L := \frac{\ln^2 a_L}{a_L^{1/4}}, \quad D = [0, L] \setminus \cup_{k=0}^i \left[ \left( z_k - \frac{3 \ln a_L}{8 \sqrt{a_L}} \right) \vee 0, \left( z_k + \frac{3 \ln a_L}{8 \sqrt{a_L}} \right) \wedge L \right].$$

- **Shape around the maximum** *The processes  $h_i$  and  $b_i$  converge to  $h$  and  $b$  uniformly over compact subsets of  $\mathbf{R}$  in probability as  $L \rightarrow \infty$ .*

**Remark 1.1** The lower bound in (4) does not hold near the zeros of the eigenfunctions and that is the reason for the restriction to  $t \in D$ . Actually our proof shows that near any zero  $z_k$  we have

$$\forall t \in \left[ \left( z_k - \frac{3 \ln a_L}{8 \sqrt{a_L}} \right) \vee 0, \left( z_k + \frac{3 \ln a_L}{8 \sqrt{a_L}} \right) \wedge L \right],$$



**Fig. 2** A very schematic plot of the fifth eigenfunction  $\varphi_5$  (not at scale). The main peak lies at  $U_5$ , while near  $U_k$ , for every  $k < 5$ , the eigenfunction has peaks of smaller order. Observe that the height of the peak near  $U_k$  decays exponentially with the distance  $|U_k - U_5|$ . Red dots correspond to the zeros of the eigenfunction (color figure online)

$$\varphi_i(t) = \varphi'_i(z_k) \frac{\sinh(\sqrt{a_L}|t - z_k|)}{\sqrt{a_L}} (1 + o(1)),$$

with a probability going to 1.

**Remark 1.2** Note that  $h$  is the main eigenfunction of the operator  $-\Delta + b'$  with  $b'$  the derivative of  $b$ .

When  $i > 1$ , we obtain further information about the eigenfunctions (see Fig. 2):

**Theorem 3** (Local maxima and zeros) Fix  $i > 1$ . Let us order the locations of the maxima of the  $i$  first eigenfunctions  $U_{\sigma(1)} < \dots < U_{\sigma(i)}$ .

- Position of the local maximum *There exists a constant  $C > 0$  such that for all  $k \in \{1, \dots, i\}$ , all the points where the eigenfunction  $|\varphi_i|$  reaches its maximum over the time interval  $[z_{k-1}, z_k]$  lie at distance at most  $C/(\sqrt{a_L} \ln a_L)$  from  $U_{\sigma(k)}$  with probability going to 1.*
- Deterministic shape around the local maxima *For all  $k \in \{1, \dots, i\}$ ,  $\varphi_i(U_{\sigma(k)} + \frac{t}{\sqrt{a_L}})/\varphi_i(U_{\sigma(k)})$  converges to  $h$  uniformly over compact subsets of  $\mathbf{R}$  in probability.*
- Zeros of the eigenfunction *For  $k \in \{1, \dots, i\}$ ,*

$$|z_k - U_{\sigma(k)} - \frac{3 \ln a_L}{4 \sqrt{a_L}}| \leq \frac{(\ln \ln a_L)^2}{\sqrt{a_L}}, \quad \text{if } U_{\sigma(k)} < U_i,$$

$$|z_{k-1} - U_{\sigma(k)} + \frac{3 \ln a_L}{4 \sqrt{a_L}}| \leq \frac{(\ln \ln a_L)^2}{\sqrt{a_L}}, \quad \text{if } U_{\sigma(k)} > U_i,$$

with probability going to 1.

**Remark 1.3** Thanks to (4), the amplitude of the local maxima on the time interval  $[z_{k-1}, z_k]$  is of order  $|\varphi_i(U_{\sigma(k)})| = |\varphi_i(U_i)| \exp((-\sqrt{a_L} + O(\kappa_L))|U_{\sigma(k)} - U_i|)$  with probability going to 1.

We now consider the case where  $\mathcal{H}_L$  is endowed with Neumann boundary conditions: we denote by  $\varphi_k^{(N)}$  and  $\lambda_k^{(N)}$  the corresponding eigenfunctions and eigenvalues. Notice that  $\varphi_k^{(N)}, \lambda_k^{(N)}$  and  $\varphi_k, \lambda_k$  are all defined on a same probability space. Our next result shows that the limiting behavior of the eigenvalues/eigenfunctions in the Neumann case is the same as in the Dirichlet case. Furthermore, the convergence can be taken jointly for the two types of boundary conditions and the limits are the same, thus showing that the choice of Dirichlet/Neumann boundary conditions does not affect the asymptotic behavior of the eigenvalues/eigenfunctions.

**Theorem 4** *For Neumann boundary conditions, the eigenvalues and eigenfunctions satisfy the same results as those stated in Theorems 1, 2 and 3. Furthermore, for every  $k \geq 1$ ,*

$$(\lambda_k^{(N)} - \lambda_k)\sqrt{a_L}, \quad \text{and} \quad (U_k^{(N)} - U_k)\sqrt{a_L},$$

converge to 0 in probability as  $L \rightarrow \infty$  and consequently the limiting random variables satisfy a.s.

$$\lambda_k^{(N),\infty} = \lambda_k^\infty, \quad \text{and} \quad U_k^{(N),\infty} = U_k^\infty.$$

**Discussion**

Before presenting the outline of the proof, let us give some motivations for studying this operator. The parabolic Anderson model (PAM) is the following Cauchy problem:

$$\partial_t u(t, x) = \partial_x^2 u(t, x) - \xi(x)u(t, x), \quad u(0, x) = \delta_0(x), \quad x \in \mathbf{R}. \quad (5)$$

One expects its solution at time  $t$  to be well approximated by the solution of the same equation restricted to a segment  $[-L/2, L/2]$ , with  $L = L(t)$  properly chosen, and endowed with Dirichlet boundary conditions. Thus, the spectral decomposition of  $\mathcal{H}_L$  would yield

$$u(t, x) \approx \sum_{k=1}^{\infty} \exp(-\lambda_k t) \varphi_k(x) \varphi_k(0). \quad (6)$$

The bottom eigenvalues and eigenfunctions of  $\mathcal{H}_L$  in the large  $L$  limit should therefore give the asymptotic behavior of the solution (6) when  $t \rightarrow \infty$ : in particular, the solution should concentrate on a few islands corresponding to the ‘‘supports’’ of the first eigenfunctions. We will provide rigorous arguments on this heuristic discussion in a future work.

If the Anderson operator is multiplied by the imaginary unit on the right hand side of (5), the equation becomes the famous Schrödinger equation which is of fundamental

importance in quantum mechanics. In this case, one needs to study the whole spectrum (not only its bottom part). This will be the object of a forthcoming work.

The discretization of  $\mathcal{H}_L$  and (5) has been investigated in many papers for a general dimension  $d$ . In this case, the Laplacian is discrete on a grid, for instance  $\mathbf{Z}^d$ , and the white noise is replaced by i.i.d random variables. The discrete operator is first defined on a finite box  $B$ , say  $B = [0, L]^d \cap \mathbf{Z}^d$ , and taken with Dirichlet boundary conditions. We refer to the book of König [20] for a state of the art on the subject. Analogous results to our Theorem 1 are known for some distributions of potentials, see for instance Biskup and König [6].

When  $d = 1$ , much more precise results are known. Let us introduce the discrete Anderson Hamiltonian

$$u \mapsto -\Delta_x u + \xi u,$$

acting on  $u : [0, L] \cap \mathbf{Z} \rightarrow \mathbf{R}$  with Dirichlet boundary conditions, and where  $\Delta_x$  denotes the discrete Laplacian and  $(\xi(x), x \in \mathbf{Z})$  are i.i.d random variables.

If the variance of  $\xi(0)$  does not depend on  $L$ , the eigenfunctions are localized [9,16,25], and the local statistics of the eigenvalues are Poissonian [23,29,30]. On the other hand, if  $\xi \equiv 0$  then the eigenfunctions are spread out and the eigenvalues are deterministic with locally regular spacings (clock-points).

The critical regime appears when the variance of  $\xi$  is of order  $1/L$ . It has been considered by Kritchevski, Valkó and Virág in [26]. They proved that delocalization holds in this case and that the eigenvalues near a fixed bulk energy  $E$  have a point process limit depending on only one parameter  $\tau$  (which is a simple function of the variance and energy  $E$ ) they called  $\text{Schr}_\tau$ . Moreover, they showed that this point process exhibits strong eigenvalue repulsion. Rifkind and Virág [34] also studied the associated eigenfunctions. They found that the shape of the eigenfunctions near their maxima is given by the exponential of a Brownian motion plus a linear drift, and is independent of the eigenvalue. Note that heuristically, this regime corresponds to the high eigenvalues and eigenfunctions of  $\mathcal{H}_L$ .

Another famous family of one-dimensional discrete random Schrödinger operators is given by the tridiagonal matrices called  $\beta$ -ensembles. These operators seen from the edge converge, in the large dimensional limit, to the stochastic Airy operator  $\mathcal{A}_\beta$  formally defined as:

$$\mathcal{A}_\beta : u \mapsto -\partial_x^2 u + \left( x + \frac{2}{\sqrt{\beta}} \xi \right) u.$$

In the paper [3], the analogue of the result of McKean was proved, namely that the first eigenvalue of the Airy operator properly rescaled converges to a Gumbel distribution in the small  $\beta$  limit. The density of states was also derived. Using similar techniques to the present paper, one should be able to prove the convergence of the point process of the first eigenvalues (properly rescaled) to a Poisson point process of intensity  $e^x dx$ . In the bulk, the eigenvalues of the  $\beta$ -ensembles converge towards the  $\text{Sine}_\beta$  process [37]. In the small  $\beta$  limit, the  $\text{Sine}_\beta$  process was also shown to converge towards a (homogeneous) Poisson point process [2] using its characterization via coupled

diffusions. Closely related discrete models are Jacobi matrices with random decaying potential [26] and CMV matrices [24].

The analogue of  $\mathcal{H}_L$  in dimension  $d$  greater than 1 can be considered. However, due to the irregularity of the white noise, the eigenvalue problem becomes singular already in dimension 2 and one has to renormalize the operator by infinite constants. This has been carried out under periodic b.c. by Allez and Chouk [1] in dimension 2, and by Gubinelli et al. [17] in dimension 3 by means of the recently introduced paracontrolled calculus [15]; the construction of the operator under Dirichlet b.c. in dimensions 2 and 3 has been performed in [27] using the theory of regularity structures [18].

Another generalization would be to consider the *multivariate* Anderson Hamiltonian of the form  $-\partial_x^2 + W'$ , operating on the vector-valued function space  $L^2([0, L], \mathbf{R}^r)$  and where  $W'$  is the derivative of a matrix valued Brownian motion. This study has been done in the case of the stochastic Airy operator by Bloemendal and Virág [8]. The eigenvalues of the multivariate Anderson Hamiltonian are characterized by a family of coupled SDEs studied by Allez and one of the authors in [4].

## Main ideas of the proof

We now present the main arguments of the proof. For the sake of clarity, we restrict ourselves to the case of Dirichlet boundary conditions.

## The Riccati transform

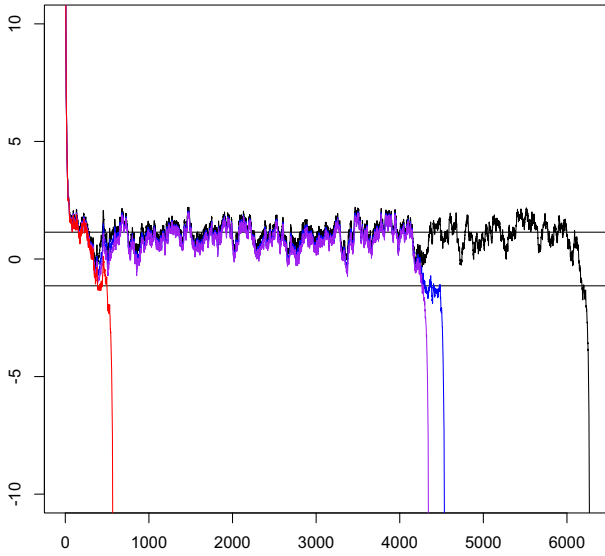
For any  $a \in \mathbf{R}$ , the differential equation  $-f'' + f\xi = -af$  on  $[0, L]$  with initial condition  $f(0) = 0$ ,  $f'(0) = 1$  admits a unique solution (by classical ODE arguments). The pair  $(-a, f)$  is then an eigenvalue/eigenfunction of the operator  $\mathcal{H}_L$  if and only if  $f(L) = 0$ . A very convenient tool for studying the solutions of this differential equation is the so-called *Riccati transform* that maps  $f$  onto  $X_a := f'/f$ . One can check that  $X_a$  starts from  $X_a(0) = +\infty$  and solves

$$dX_a(t) = (a - X_a(t)^2)dt + dB(t), \quad t \in [0, L], \quad (7)$$

where  $B(t) := \langle \xi, \mathbf{1}_{[0,t]} \rangle$  is a one-dimensional Brownian motion. Note that, whenever  $X_a$  hits  $-\infty$  (that is, whenever  $f$  vanishes) it is restarted from  $+\infty$ . The set of eigenvalues is then in one-to-one correspondence with the set of points  $-a$  such that  $X_a(L) = -\infty$ .

The key point of the Riccati transform is that the processes  $X_a$ 's satisfy the following *monotonicity property*: if  $a < a'$  then  $X_a(\cdot) \leq X_{a'}(\cdot)$  up to the first hitting time of  $-\infty$  of  $X_a$ , see Fig. 3. As a consequence, for any  $k \geq 1$ , the r.v.  $-\lambda_k$  is the largest  $a \in \mathbf{R}$  such that  $X_a$  explodes exactly  $k$  times on  $[0, L]$ : the Riccati transform  $\mathcal{X}_k = \varphi'_k/\varphi_k$  of the  $k$ -th eigenfunction  $\varphi_k$  is then given by  $X_{-\lambda_k}$ . (Note that the above discussion relies on deterministic arguments so that the characterization of the  $\lambda_k, \varphi_k$ 's in terms of the processes  $X_a$ 's holds true almost surely.)





**Fig. 3** Ordering of the diffusions  $X_a$  until their first explosion time for  $a = 0, 55$  (red),  $a = 0, 7$  (purple),  $a = 0, 845$  (blue),  $a = 1, 3$  (black) (color figure online)

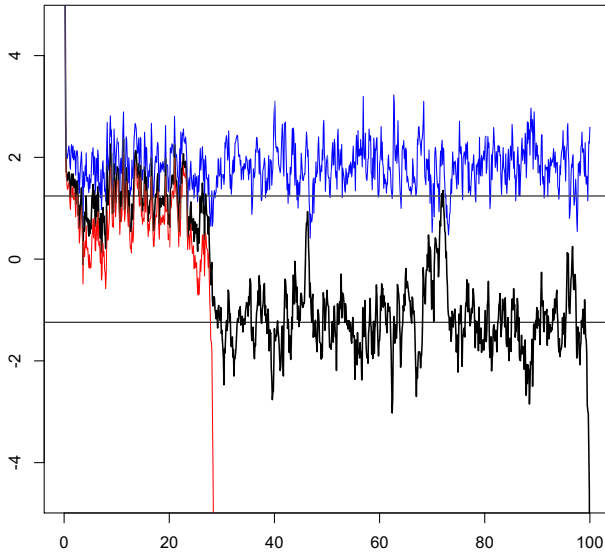
The study of the eigenvalues/eigenfunctions thus boils down to the study of the collection of coupled processes  $X_a$ 's. For any given  $a \in \mathbf{R}$ , the process  $X_a$  is an irreversible diffusion that evolves in the potential  $V_a$ :

$$V_a(x) = \frac{x^3}{3} - ax, \quad \text{for all } x \in \mathbf{R}, \quad a \in \mathbf{R}, \tag{8}$$

which admits a well whenever  $a > 0$  (see Fig. 7). For any  $a \in \mathbf{R}$ , this diffusion hits  $-\infty$  in finite time almost surely. The aforementioned result of McKean shows that, as  $L$  becomes large, the first eigenvalue goes to  $-\infty$  at speed  $a_L$  (and the same holds for the next  $k$  eigenvalues). Therefore, to study the bottom of the spectrum of  $\mathcal{H}_L$  we can focus on diffusions with a large parameter  $a > 0$ .

A typical realization of the diffusion  $X_a$  for a fixed large  $a > 0$  does the following. It comes down from  $+\infty$  very quickly and oscillates around  $\sqrt{a}$  (bottom of the well of  $V_a$ ) for a long time. It makes many attempts to get out of the well, and from time to time, it makes an exceptional excursion to  $-\sqrt{a}$ , spends a very short time near that point, and either goes back to  $\sqrt{a}$  or explodes to  $-\infty$  very quickly.

Since the eigenvalues  $(\lambda_k, k \in \mathbf{N})$  depend on the realization of the underlying noise  $\xi$ , the process  $\chi_k$  is *not* a diffusion and actually does not look like a typical solution of (7). For instance, we will see later that  $\chi_k$  spends a macroscopic time near the unstable point  $-\sqrt{|\lambda_k|}$ , see Figs. 4 and 5. Nevertheless, we will see that it is possible to extract information out of the realizations of *typical* diffusions  $X_a$ 's.



**Fig. 4** Simulation of  $\chi_1$  (in black) and of two diffusions  $X_a$  (in red) and  $X_{a'}$  (in blue) with  $a < -\lambda_1 < a'$  (color figure online)

**Convergence of the eigenvalues**

Let us present the main steps of the proof of our theorems. We start with the convergence of the point process of the rescaled and recentered eigenvalues

$$Q_L := \sum_{k \geq 1} \delta_{4\sqrt{a_L}(\lambda_k + a_L)}$$

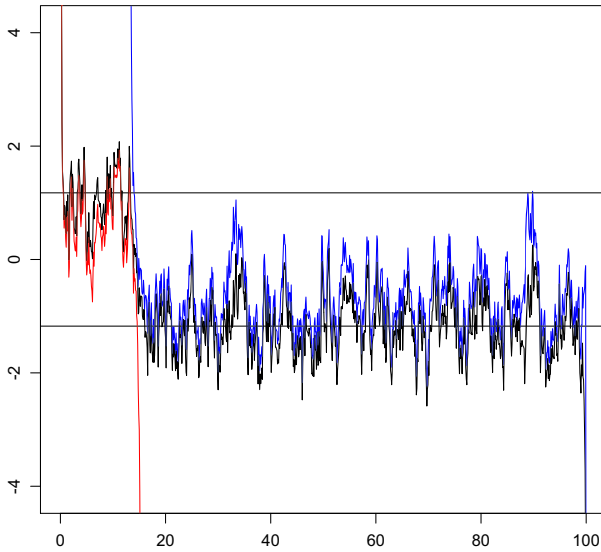
of Theorem 1. While tightness is easy to get, we identify the limit by showing that for all fixed  $r_1 < \dots < r_k$  the random variable  $(Q_L((r_1, r_2]), \dots, Q_L((r_{k-1}, r_k]))$  converges to a vector of independent Poisson r.v. with parameters  $e^{r_i} - e^{r_{i-1}}$ .

Observe that  $Q_L((r_{i-1}, r_i]) = \#X_{a_i} - \#X_{a_{i-1}}$  where  $a_i = \sqrt{a_L} - r_i / (4\sqrt{a_L})$ . We subdivide  $(0, L]$  into the  $2^n$  disjoint subintervals  $(t_j^n, t_{j+1}^n]$  with  $t_j^n = j2^{-n}L$ , and we introduce the diffusion  $X_{a_i}^j$  that starts from  $+\infty$  at time  $t_j^n$  and follows the SDE (7) with parameter  $a_i$ . Then, we show that with large probability for all  $i$  and  $j$ :

- $X_{a_i}$  explodes at most one time on  $(t_j^n, t_{j+1}^n]$ ,
- $X_{a_i}$  explodes on  $(t_j^n, t_{j+1}^n]$  if and only if  $X_{a_i}^j$  explodes on  $(t_j^n, t_{j+1}^n]$ .

This holds because the potential  $V_{a_i}$  possesses a large well and the diffusion goes down from  $+\infty$  into the well in a very short time: its starting point quickly becomes irrelevant.

Therefore, it suffices to deal with the diffusions  $X_{a_i}^j$  restricted to  $(t_j^n, t_{j+1}^n]$ : these diffusions are independent for different values of  $j$ , and are monotone in  $i$ , so that a very simple computation, see Lemma 2.7, allows to get the aforementioned convergence.



**Fig. 5** Simulation of  $\chi_1$  (in black) with the two diffusions  $X_a$  (in red) and  $\hat{X}_a$  (in blue) until their first explosion, with  $a < -\lambda_1$  (color figure online)

An important remark is that we only need the *monotonicity* of the coupled diffusions for this part of the proof: we use no finer information about this coupling. The arguments in the proof could be applied to other situations where the number of explosions of coupled diffusions counts the eigenvalues (e.g. Airy $_{\beta}$  [33], Sine $_{\beta}$  [37] or the Stochastic Bessel Operator [32]).

**The first eigenfunction**

Let us now concentrate on the asymptotic behavior of the first eigenfunction. According to the convergence (2), we consider a discretization  $M_L$  of a small neighborhood of  $a_L$  of mesh  $\varepsilon/\sqrt{a_L}$  for a fixed small  $\varepsilon > 0$  and we will use precise estimates on the typical behavior of  $X_a$ , simultaneously for all  $a \in M_L$ . With large probability,  $-\lambda_1$  falls within this neighborhood of  $a_L$  and there exist two points  $a, a'$  in the discretization such that  $-\lambda_2 < a \leq -\lambda_1 < a'$ . By definition,  $X_a$  explodes one time and  $X_{a'}$  does not explode on the time interval  $[0, L]$ .

From the monotonicity property, we have  $X_a \leq \chi_1 < X_{a'}$  up to the first explosion time of  $X_a$ . With large probability,  $X_a$  and  $X_{a'}$  are “typical”: in particular, they remain close to  $\sqrt{a}$  and  $\sqrt{a'}$  respectively most of the time. If the mesh of the discretization  $M_L$  has been chosen small enough, we deduce that  $\chi_1$  is squeezed in between those two typical diffusions that are close to each other with large probability, up to the first explosion time of  $X_a$ . However, this does not provide any good control on  $\chi_1$  after this explosion time. In particular, it does not say whether, for instance,  $\chi_1$  remains around  $-\sqrt{|\lambda_1|}$ , or goes back to  $\sqrt{|\lambda_1|}$ . To push the analysis further, we rely on a symmetry argument.

The Riccati transform can be applied not only forward in time from time 0 but also *backward in time* from time  $L$ . This yields another set of diffusions  $(\hat{X}_a, a \in \mathbf{R})$ , called the *time-reversed diffusions*, which is equal in law to  $(-X_a, a \in \mathbf{R})$ . Namely, we have:

$$d\hat{X}_a(t) = V'_a(\hat{X}_a(t))dt + d\hat{B}(t),$$

starting from  $\hat{X}_a(0) = -\infty$  and where  $\hat{B}(t) := B(L-t) - B(L)$ . A coupling argument shows that the number of explosions of  $X_a$  and  $\hat{X}_a$  coincide. This allows to say that  $\chi_1$ , run backward in time from time  $L$ , is squeezed in between  $\hat{X}_a$  and  $\hat{X}_{a'}$  up to the explosion time of  $\hat{X}_a$ . Then, it remains to show that the explosion times of  $X_a$  and  $\hat{X}_a$  are very close so that the intervals on which we bound the first eigenfunction overlap (see Fig. 5).

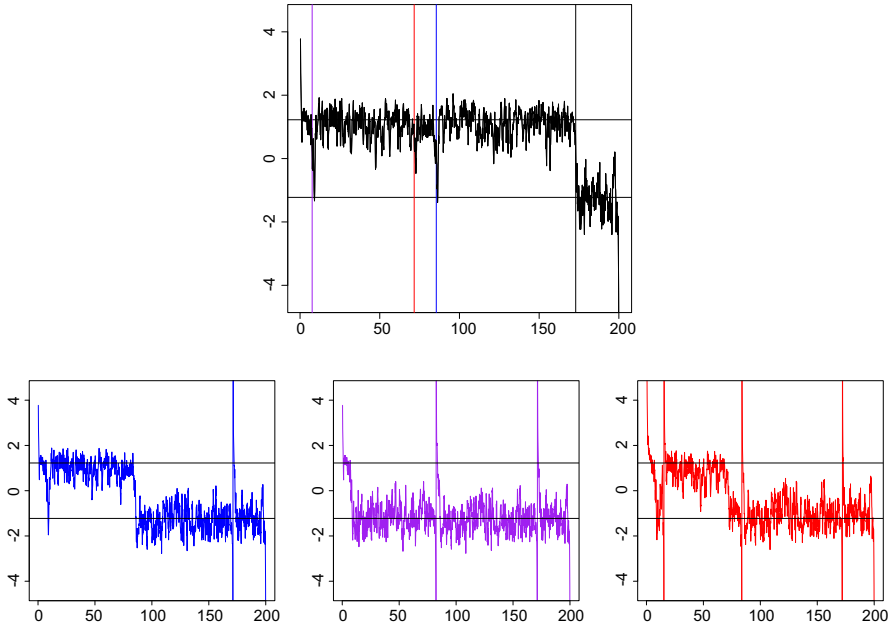
Since  $X_a$  and  $X_{a'}$  spend most of their time near  $\sqrt{a} \simeq \sqrt{a'} \simeq \sqrt{a_L}$ , we deduce by inverting the Riccati transform, that the eigenfunction grows exponentially fast at this rate up to the explosion time of  $X_a$ . Then, it follows the time-reversed diffusions which spend most of their time near  $-\sqrt{a_L}$  so that the first eigenfunction decays exponentially fast from the explosion time until time  $L$ . This proves the localization of the first eigenfunction near the explosion time of  $X_a$ .

We push this analysis further and obtain a much more precise result on the localization, namely the *shape of the first eigenfunction* of Theorem 2. The explosion of  $X_a$  occurs right after an exceptional descent from  $+\sqrt{a}$  to  $-\sqrt{a}$  whose trajectory is concentrated around the deterministic function  $t \mapsto -\sqrt{a} \tanh(\sqrt{a} t)$  (appropriately shifted in time). This is a simple consequence of the large deviations principle satisfied by the diffusion, that ensures that the trajectory on this exceptional descent is roughly given by the solution of the ODE

$$dx(t) = V'_a(x(t))dt.$$

We will obtain a precise statement about this fact thanks to the Girsanov Theorem. We then show that  $\chi_1$  remains very close to  $X_a$  upon this deterministic descent. A careful argument, involving the time-reversed diffusion, shows that the maximum of  $\varphi_1$  is located at one of the zeros of  $\chi_1$  achieved during this deterministic descent: since all these zeros lie in a tiny region, we get a precise enough control on the location of the maximum. Applying the inverse of the Riccati transform to  $-\sqrt{a} \tanh(\sqrt{a} t)$ , one gets the deterministic shape  $h(t)$  of the statement of the theorem.

Up to this point, we have not proved yet that the localization center is asymptotically uniform on  $[0, L]$ . To that end, we present a coupling argument that relates the localization centers of the first eigenfunctions of  $\mathcal{H}_L$  and of the operator  $\tilde{\mathcal{H}}_L$ , the latter being obtained upon replacing the white noise by its image through a Lebesgue preserving bijection. This coupling argument eventually shows that the law of these centers is invariant under (a large enough class of) Lebesgue preserving bijections so that it is necessarily uniform.



**Fig. 6** Simulation of the Riccati transform of the first 4 eigenfunctions (successively black, blue, purple, red). On the first picture, the position of the maxima of the first four eigenfunctions are represented with their color. Notice that the number of explosions equals the numbering of the eigenfunction. The explosions occur near the maxima of the previous eigenfunctions (color figure online)

**The next eigenfunctions**

For the  $k$ -th eigenfunction with  $k > 1$ , the situation is slightly more involved though the strategy is the same. We show that with large probability,  $-\lambda_k$  falls in the aforementioned neighborhood of  $a_L$  and that there exist  $a, a'$  in the discretization of that neighborhood such that  $-\lambda_{k+1} < a \leq -\lambda_k < a' \leq -\lambda_{k-1}$ . The typical diffusions  $X_a$  and  $X_{a'}$  explode  $k$  and  $k - 1$  times respectively. We show that they remain close to each other up to the *additional* explosion time of  $X_a$ , thus providing a good control on  $\mathcal{X}_k$  up to this time. Then, we rely on the time-reversed diffusions to complete the picture, as in the case of the first eigenfunction. We refer to Fig. 6 for an illustration.

**Organization of the paper**

In Sect. 2, we collect some first estimates on the diffusion  $X_a$  and we prove the convergence of the point process of eigenvalues towards the Poisson point process with intensity  $e^x dx$ . The proofs of the first estimates are postponed to Sect. 4. In Sect. 3, we prove Theorems 1, 2 and 3. The proofs are based on more precise estimates on the diffusion  $X_a$  on its exceptional excursions to  $-\sqrt{a}$ , which are obtained

in Sect. 5. Finally, in Sect. 6 we prove Theorem 4 about Neumann boundary conditions.

## Notations

The first hitting time by a continuous function  $f$  of a point  $\ell \in \mathbf{R} \cup \{-\infty\}$  is defined as follows

$$\tau_\ell(f) := \inf\{t \geq 0 : f(t) = \ell\},$$

Moreover, for any integrable function  $f$  we define its average on  $[s, t]$  by setting

$$\bar{f}_s^t := \frac{1}{t-s} \int_s^t f(x) dx.$$

When  $s = t$ , we set this quantity to  $f(s)$ .

## 2 Convergence of the eigenvalues towards the Poisson point process

The goal of this section is to prove that the point process  $\mathcal{Q}_L := (4\sqrt{a_L}(\lambda_k + a_L))$  of the rescaled eigenvalues converges in law to a Poisson point process of intensity  $e^x dx$ . This is the first step in the proof of Theorem 1. After some technical preliminaries on the well-definiteness of our diffusions, we collect some first estimates on the diffusion  $X_a$ , and then, we prove the convergence.

### 2.1 Technical preliminaries

For any given continuous function  $b$  starting from 0, we consider the ODE:

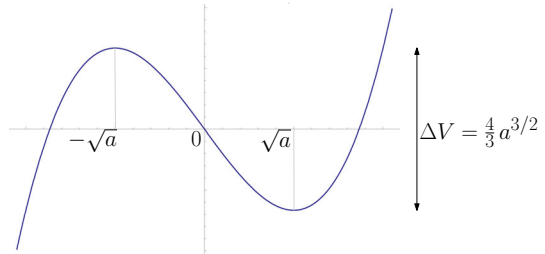
$$\begin{cases} x(t) = -\int_0^t V'_a(x(s)) ds + b(t), & t > 0, \\ x(0) = +\infty. \end{cases}$$

This ODE admits a unique solution that leaves  $+\infty$  at time  $0+$ , and restarts from  $+\infty$  whenever it hits  $-\infty$ . Let  $B$  be a standard Brownian motion on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . Without loss of generality, we can assume that  $B(\omega)$  is continuous for all  $\omega \in \Omega$ . We apply deterministically the solution map associated to the ODE above for every realization of the Brownian motion. This yields a well-defined process  $(X_a(t), t \geq 0, a \in \mathbf{R})$ .

**Remark 2.1** Actually, the process can be initialized not only at time 0 but also at some arbitrary time  $t_0 \geq 0$ . The above construction still applies, and yields a family  $(X_a^{t_0}(t), t \geq t_0, a \in \mathbf{R}, t_0 \geq 0)$  with  $X_a^{t_0}(t_0) = +\infty$ .

The monotonicity of the solution map associated to the ODE yields the following property: if  $a < a'$  and  $X_a(t) \leq X_{a'}(t)$  then  $X_a(t + \cdot) \leq X_{a'}(t + \cdot)$  up to the next explosion time of  $X_a$ .

**Fig. 7** The potential  $V_a$  when  $a > 0$



**2.2 First estimates on the diffusion  $X_a$**

We now study the process just defined above:

$$dX_a(t) = (a - X_a(t)^2)dt + dB(t), \quad X_a(0) = +\infty.$$

This diffusion evolves in the potential  $V_a(x) = -ax + x^3/3$ . It blows-up to  $-\infty$  in finite time a.s. and immediately restarts from  $+\infty$ . From now on, we let  $0 = \zeta_a(0) < \zeta_a(1) < \zeta_a(2) < \dots$  be the successive explosion times of the diffusion  $X_a$ . When  $a > 0$ , the potential  $V_a$  has a well centered at  $\sqrt{a}$ . The diffusion has to cross the barrier from  $\sqrt{a}$  to  $-\sqrt{a}$ , which is of size  $\Delta V_a = (4/3)a^{3/2}$ , in order to explode to  $-\infty$  (see Fig. 7).

The expectation of the first explosion time  $\zeta_a(1)$  admits the following expression, see [14]:

$$m(a) := \mathbb{E}[\zeta_a(1)] = \sqrt{2\pi} \int_0^{+\infty} \frac{dv}{\sqrt{v}} \exp\left(2av - \frac{1}{6}v^3\right) \tag{9}$$

$$= \frac{\pi}{\sqrt{a}} \exp\left(\frac{8}{3}a^{3/2}\right) (1 + o(1)), \quad \text{when } a \rightarrow +\infty. \tag{10}$$

We see that the explosion time is of order  $\exp(2\Delta V_a)$ , which is in line with Kramers' law.

**Remark 2.2** Recall that the number of explosions of  $X_a$  in  $[0, L]$  is equal to the number of eigenvalues below  $-a$ . By the law of large numbers, it readily implies that the integrated density of states  $N(\lambda)$  mentioned in the introduction is precisely given by

$$N(\lambda) = \frac{1}{m(-\lambda)} \quad \text{for all } \lambda \in \mathbf{R}. \tag{11}$$

The following result is due to McKean, we refer to [3,28] for a proof.

**Proposition 2.3** *As  $a \rightarrow \infty$ , the r.v.  $\zeta_a(1)/m(a)$  converges in distribution to an exponential law of parameter 1.*

As a direct corollary, we get the following result.

**Proposition 2.4** *The point process of the explosion times of  $X_a$  rescaled by  $m(a)$  converges in law, for the vague topology, to a Poisson point process on  $\mathbf{R}_+$  with intensity 1.*

It is then natural to introduce  $L \mapsto a_L$  as the reciprocal of  $a \mapsto m(a)$ . A simple computation yields the following asymptotics

$$a_L = \left(\frac{3}{8} \ln L\right)^{2/3} + 3^{-1/3} \cdot 2^{-1} (\ln L)^{-1/3} \left(\ln\left(\frac{3^{1/3}}{2\pi} \ln^{1/3} L\right) + o(1)\right). \tag{12}$$

The diffusion  $X_a$ , with  $a$  close to  $a_L$ , typically explodes a finite number of times on the time interval  $[0, L]$ . More precisely, for all  $r \in \mathbf{R}$ , using (10), we have as  $L \rightarrow \infty$ :

$$m\left(a_L + \frac{r}{\sqrt{a_L}}\right) = \frac{\pi}{\sqrt{a_L}} \exp\left(\frac{8}{3} a_L^{3/2} + 4r\right) (1 + o(1)),$$

so that the order of magnitude of the number of explosions of  $X_a$  on  $[0, L]$  for  $a = a_L + r/(4\sqrt{a_L})$  is given by  $L/m(a_L + r/(4\sqrt{a_L})) = \exp(-r)(1 + o(1))$ .

Until the end of this subsection, we drop the subscript  $a$  from  $X_a$  and from the explosion times  $\zeta_a(k)$ ,  $k \geq 0$  to alleviate the notations. Let us analyse the diffusion  $X$  until its first explosion time  $\zeta := \zeta(1)$ . We set

$$t_a = \frac{\ln a}{\sqrt{a}}, \quad \text{and} \quad h_a = \frac{\ln a}{a^{1/4}}. \tag{13}$$

A typical realization of the process  $X$  is close to a deterministic path when it comes down from  $+\infty$  (entrance) and when it explodes to  $-\infty$  (explosion):

**1- Entrance** Let  $\ell := \sqrt{a} + h_a$ . We have

$$\tau_\ell(X) = \frac{3}{8} t_a - \frac{\ln \ln a}{2\sqrt{a}} + O\left(\frac{1}{\sqrt{a}}\right),$$

together with the bound

$$\sup_{t \in (0, \frac{3}{8} t_a]} |X(t) - \sqrt{a} \coth(\sqrt{a}t)| \leq 1.$$

**2- Explosion** Let  $\ell := -\sqrt{a} - h_a$ . We have

$$\tau_{-\infty}(X) - \tau_\ell(X) = \frac{3}{8} t_a - \frac{\ln \ln a}{2\sqrt{a}} + O\left(\frac{1}{\sqrt{a}}\right).$$

together with the bound

$$\sup_{t \in [\tau_\ell(X), \zeta]} |X(t) + \sqrt{a} \coth(\sqrt{a}(\zeta - t))| \leq 1.$$



We denote by  $\mathcal{D}_a^N$  the event on which the  $N$  first trajectories  $(X(t), t \in [\zeta(k), \zeta(k+1)))$  for  $k \in \{0, \dots, N - 1\}$  satisfy the estimates above.

**Proposition 2.5** (Entrance and explosion) *Fix  $N \in \mathbb{N}$ . For all the parameters  $a$  large enough, we have*

$$\mathbb{P}[\mathcal{D}_a^N] \geq 1 - \frac{1}{\ln a}.$$

The proof of Proposition 2.5 relies on a comparison with an ordinary differential equation (in the same way as in [12]): it is postponed to Sect. 4.

For  $t_0 > 0$ , let  $X^{t_0}$  be the diffusion following the same SDE as  $X$  but starting from  $+\infty$  at time  $t_0$ . The following *synchronization* estimate ensures that the diffusion  $X^{t_0}$  gets very close to the original diffusion  $X$  after a short time of order  $O(t_a)$ , and that they stay together until their first explosion.

We set  $\tau_{-2\sqrt{a}}^k(X) := \inf\{t \geq \zeta_a(k - 1) : X(t) = -2\sqrt{a}\}$ .

**Proposition 2.6** (Synchronization) *Fix  $N \geq 1$  and  $(t_0(a))_{a>0}$  a family of deterministic starting times which may depend on  $a$ . There exists  $C > 0$  such that the following holds true with a probability at least  $1 - O(1/\ln a)$  as  $a \rightarrow \infty$ . If there exists  $k \in \{1, \dots, N\}$  such that  $t_0(a) \in [\zeta_a(k - 1), \zeta_a(k))$  then we have the bounds*

$$\begin{aligned} |X(t) - X^{t_0(a)}(t)| &\leq h_a, \quad \forall t \in [t_0(a) + \frac{3}{8}t_a, \tau_{-2\sqrt{a}}^k(X)], \\ |\tau_{-\infty}(X^{t_0(a)}) - \zeta_a(k)| &< \frac{C}{\sqrt{a}}. \end{aligned}$$

The proof of Proposition 2.6 is based on a coupling with a stationary diffusion. Since the arguments are rather standard, we postpone the proof to Sect. 4.

### 2.3 Proof of the convergence of the point process of the eigenvalues

We consider

$$\mathcal{Q}_L := \sum_{k \geq 1} \delta_{4\sqrt{a_L}(\lambda_k + a_L)},$$

which is a random element of the set of measures on  $\mathbf{R}$  with finite mass on  $(-\infty, r)$  for any  $r \in \mathbf{R}$ , and endowed with the topology that makes continuous the maps  $\mu \mapsto \int f \mu$  for all bounded and continuous functions  $f$  with support bounded on the right. By [21, Lemma 14.15], the family  $(\mathcal{Q}_L)_{L>1}$  is tight if for every  $r > 0$  and every  $\varepsilon > 0$ , there exists  $c > 0$  such that

$$\sup_{L>1} \mathbb{P}[\mathcal{Q}_L((-\infty, r]) > c] < \varepsilon. \tag{14}$$

Since  $\mathcal{Q}_L((-\infty, r])$  is equal to the number of explosions of  $X_{a_L-r/(4\sqrt{a_L})}$  on the time-interval  $[0, L]$ , we deduce from Proposition 2.4 that there exists  $c' > 0$  such that

$$\overline{\lim}_{L \in \mathbb{N}, L \rightarrow \infty} \mathbb{P}[\mathcal{Q}_L((-\infty, r]) > c'] < \varepsilon,$$

and therefore, that there exist  $c > 0$  such that

$$\sup_{L \in \mathbb{N}} \mathbb{P}[\mathcal{Q}_L((-\infty, r]) > c] < \varepsilon.$$

Since  $a_L = a_{\lfloor L \rfloor} + o(1/\sqrt{a_{\lfloor L \rfloor}})$ , we can bound the number of explosions of  $X_{a_L-r/(4\sqrt{a_L})}$  by the number of explosions of  $X_{a_{\lfloor L \rfloor}-r'/(4\sqrt{a_{\lfloor L \rfloor}})}$  for some appropriately chosen  $r'$ , thus ensuring that the latter bound extends to all  $L \in (1, \infty)$ . We deduce that (14) holds true, thus ensuring the tightness of  $(\mathcal{Q}_L)_{L>1}$ .

It remains to prove that any limit  $\mathcal{Q}_\infty$  is distributed as a Poisson point process with intensity  $e^x dx$ . By classical arguments, it suffices to show that for every  $k \geq 1$ , and every  $-\infty = r_0 < r_1 < r_2 < \dots < r_k$ , if we set

$$\mathcal{Q}_L(i) := \mathcal{Q}_L((r_{i-1}, r_i]), \quad i = 1, \dots, k,$$

then  $(\mathcal{Q}_L(1), \dots, \mathcal{Q}_L(k))$  converges in law to a vector of independent Poisson r.v. with parameters  $p_i$  where

$$p_i := e^{r_i} - e^{r_{i-1}}, \quad i = 1, \dots, k.$$

Let us fix from now on  $r_1 < r_2 < \dots < r_k$ , and notice that if we denote by  $\#X_{a_i}$  the number of explosions of  $X_{a_i}$  on  $[0, L]$ , where  $a_i = a_L - r_i/(4\sqrt{a_L})$ , then we have almost surely

$$\mathcal{Q}_L(i) = \#X_{a_i} - \#X_{a_{i-1}}, \quad i = 1, \dots, k.$$

To identify the limiting law of  $(\mathcal{Q}_L(1), \dots, \mathcal{Q}_L(k))$ , we follow an indirect approach. First, we introduce some simpler r.v.  $\tilde{\mathcal{Q}}_L(i)$  which are shown to converge to the right limits, and then we prove that they are actually close in probability to the original ones.

We discretize the time-interval  $[0, L]$  in such a way that the studied diffusions will typically explode at most once on each sub-interval with large probability. Let  $n \geq 1$  be given. We set  $t_j^n := j2^{-n}L$  for all  $j \in \{0, \dots, 2^n\}$  and we let  $X_a^j := X_a^{t_j^n}$  be the diffusion starting from  $+\infty$  at time  $t_j^n$ . The main idea of the proof is to approximate the number of explosions of  $X_{a_i}$  on  $(t_j^n, t_{j+1}^n]$  by the number of explosions of  $X_{a_i}^j$  on the same interval. Such an approximation is justified by Proposition 2.6. The advantage of considering the diffusions  $X_{a_i}^j$  restricted to  $(t_j^n, t_{j+1}^n]$  is twofold: first, for every  $j$ , they are ordered in  $i$  by the monotonicity of the diffusions; second, for different values  $j$  they are independent.

For every  $j \in \{0, \dots, 2^n - 1\}$ , and for every  $i \in \{1, \dots, k\}$ , we set  $U_j(i) = 1$  if  $X_{a_i}^j$  explodes on  $(t_j^n, t_{j+1}^n]$ , and  $U_j(i) = 0$  otherwise. Recall that  $a_1 > a_2 > \dots > a_k$ . Note that on each time-interval  $(t_j^n, t_{j+1}^n]$ , we have the ordering  $X_{a_k}^j \leq \dots \leq X_{a_1}^j$  until the first explosion of  $X_{a_k}^j$ .

Then, we set

$$\tilde{Q}_L^{(n)}(i) := \sum_{j=0}^{2^n-1} (U_j(i) - U_j(i-1)), \quad i = 1, \dots, k.$$

**Lemma 2.7** *The vector  $(\tilde{Q}_L^{(n)}(i))_{i=1, \dots, k}$  converges in distribution, as  $L \rightarrow \infty$  and then  $n \rightarrow \infty$ , to a vector of independent Poisson r.v. with parameters  $p_i$ .*

**Proof** First of all, the collection (indexed by  $j = 0, \dots, 2^n - 1$ ) of the  $k$ -dimensional vectors  $(U_j(i), i = 1, \dots, k)$  is i.i.d., and for every  $j$  we have almost surely  $U_j(1) \leq \dots \leq U_j(k)$  since  $a_1 > \dots > a_k$ . Henceforth we have for  $i \in \{1, \dots, k - 1\}$ ,

$$\mathbb{P}[U_j(1) = \dots = U_j(i) = 0, U_j(i+1) = \dots = U_j(k) = 1] = \mathbb{P}[U_j(i) = 0, U_j(i+1) = 1] = \mathbb{P}[U_j(i+1) = 1] - \mathbb{P}[U_j(i) = 1],$$

as well as

$$\mathbb{P}[U_j(1) = \dots = U_j(k) = 0] = \mathbb{P}[U_j(k) = 0],$$

and

$$\mathbb{P}[U_j(1) = \dots = U_j(k) = 1] = \mathbb{P}[U_j(1) = 1].$$

Consequently, the law of the vector  $U_j$  is completely characterized by its one-dimensional marginals. Since  $m(a_L)/m(a_i) \rightarrow e^{r_i}$  as  $L \rightarrow \infty$ , Proposition 2.4 gives as  $L \rightarrow \infty$

$$\mathbb{P}[U_j(i) = 1] \rightarrow 1 - \exp(-2^{-n} e^{r_i}),$$

for any  $j$ . Moreover, the random variables  $U_j, j \in \{0, \dots, 2^n - 1\}$  are independent. We have all the elements at hand to compute the limiting distributions of the  $\tilde{Q}_L^{(n)}(i)$ 's. Fix  $\ell_1, \dots, \ell_k$  such that  $\ell = \sum_1^k \ell_i \leq 2^n$ . Then, we find:

$$\begin{aligned} & \lim_{L \rightarrow \infty} \mathbb{P}\left[\bigcap_{i=1}^k \{\tilde{Q}_L^{(n)}(i) = \ell_i\}\right] \\ &= \lim_{L \rightarrow \infty} \binom{2^n}{\ell_1, \dots, \ell_k, 2^n - \ell} \end{aligned}$$

$$\begin{aligned}
 & \left( \prod_{i=1}^k \mathbb{P}[U_1(1) = \dots = U_1(i-1) = 0, U_1(i) = \dots = U_1(k) = 1]^{\ell_i} \right) \\
 & \times \mathbb{P}[U_1(1) = \dots = U_1(k) = 0]^{2^n - \ell} \\
 & = \binom{2^n}{\ell_1, \dots, \ell_k, 2^n - \ell} \\
 & \left( \prod_{i=1}^k (1 - \exp(-2^{-n} e^{r_i}) - 1 + \exp(-2^{-n} e^{r_{i-1}}))^{\ell_i} \right) \exp(-2^{-n} e^{r_k} (2^n - \ell)) \\
 & = \prod_{i=1}^k \frac{(e^{r_i} - e^{r_{i-1}})^{\ell_i}}{\ell_i!} e^{-(e^{r_i} - e^{r_{i-1}})} (1 + O(2^{-n})),
 \end{aligned}$$

so that the limiting distribution as  $n \rightarrow \infty$  is the product of Poisson distributions with parameter  $p_i$ . □

We now introduce an event on which  $\tilde{Q}_L^{(n)}$  and  $Q_L$  coincide. Let  $\mathcal{F}_{L,n}$  be the event on which for every  $i \in \{1, \dots, k\}$  and  $j \in \{0, \dots, 2^n - 1\}$ :

- (i)  $X_{a_i}$  never explodes twice in any interval  $(t_j^n, t_{j+1}^n]$ ,
- (ii)  $X_{a_i}$  explodes on  $(t_j^n, t_{j+1}^n]$  if and only if  $X_{a_i}^j$  explodes on  $(t_j^n, t_{j+1}^n]$ .

On the event  $\mathcal{F}_{L,n}$ , the auxiliary random variables  $\tilde{Q}_L^{(n)}(i)$  indeed coincide with  $Q_L(i)$ .

If we prove that the probability of  $\mathcal{F}_{L,n}$  tends to 1 when  $L \rightarrow \infty$  and then  $n \rightarrow \infty$ , then by Lemma 2.7 we get:

$$\mathbb{P}[Q_L(1) = \ell_1, \dots, Q_L(k) = \ell_k] = \prod_{i=1}^k \frac{p_i^{\ell_i}}{\ell_i!} e^{-p_i} + \varepsilon_{L,n},$$

where  $\varepsilon_{L,n}$  goes to 0 as  $L$  and then  $n$  go to infinity. This would ensure that the vector  $Q_L$  converges in distribution to a vector of independent Poisson random variables with parameters  $p_i$ , and would conclude the proof of this section.

Therefore, it remains to show that  $\mathbb{P}[\mathcal{F}_{L,n}]$  goes to 1 as  $L \rightarrow \infty$  and then  $n \rightarrow \infty$ . By Propositions 2.4 and 2.6, the probability that for all  $i \in \{1, \dots, k\}$ :

- (a)  $X_{a_i}$  never explodes twice in any interval  $(t_j^n, t_{j+1}^n]$ ,
- (b)  $X_{a_i}$  never explodes in any interval  $(t_{j+1}^n - t_{a_L}, t_{j+1}^n]$ ,
- (c) the next explosion times of  $X_{a_i}$  and  $X_{a_i}^j$  after time  $t_j^n$  lie within a distance of order  $1/\sqrt{a_L}$  from each other,

goes to 1 as  $L \rightarrow \infty$  and then  $n \rightarrow \infty$ . Property (a) ensures (i). The monotonicity of the diffusions ensures the following assertion: if  $X_{a_i}^j$  explodes on  $(t_j^n, t_{j+1}^n]$ , then  $X_{a_i}$  explodes on  $(t_j^n, t_{j+1}^n]$ . The converse assertion is implied by (b) and (c), so that (ii) follows.

### 3 Localization of the eigenfunctions

In this section, we first introduce the time-reversed diffusions associated to the eigenvalue problem. Then, we collect some fine estimates on the diffusion during its exceptional excursion that leads it from  $\sqrt{a}$  to  $-\sqrt{a}$ . Finally, we provide the proof of Theorems 1, 2 and 3.

#### 3.1 Time reversal

We define the time-reversed operator

$$\hat{\mathcal{H}}_L = -\partial_x^2 + \hat{\xi}, \quad x \in (0, L),$$

with homogeneous Dirichlet boundary conditions, where  $\hat{\xi}(\cdot) := \xi(L - \cdot)$  in the distributional sense. Let  $(\hat{\varphi}_k, \hat{\lambda}_k)_{k \geq 1}$  be the corresponding eigenfunctions/eigenvalues and observe that almost surely for all  $k \geq 1$

$$\hat{\varphi}_k(\cdot) = \varphi_k(L - \cdot), \quad \hat{\lambda}_k = \lambda_k.$$

We can therefore study the Riccati transforms of the reversed operator and introduce the associated family of diffusions. For convenience, we take the opposite of the Riccati transform:

$$\hat{\chi}_k(t) := -\frac{\hat{\varphi}'_k(t)}{\hat{\varphi}_k(t)}, \quad t \in (0, L),$$

with boundary conditions  $\hat{\chi}_k(0) = -\infty$  and  $\hat{\chi}_k(L) = +\infty$ . We then have the identity:

$$\hat{\chi}_k(L - t) = \chi_k(t).$$

It is natural to introduce the collection of diffusions  $(\hat{X}_a(t), t \geq 0), a \in \mathbf{R}$ , as solutions of:

$$d\hat{X}_a(t) = V'_a(\hat{X}_a(t))dt + d\hat{B}(t), \tag{15}$$

starting from  $\hat{X}_a(0) = -\infty$  and where  $\hat{B}(t) := B(L - t) - B(L)$ . We then let  $\hat{\zeta}_a(k)$  be the  $k$ -th explosion time of  $\hat{X}_a$ . We also set  $\hat{\zeta}_a(0) = \zeta_a(0) = 0$ .

**Remark 3.1** Time 0 for the diffusion  $\hat{X}_a$  corresponds to time  $L$  for the diffusion  $X_a$ .

**Lemma 3.2** *Almost surely, for every  $k \geq 1$  if  $X_a$  explodes to  $-\infty$  exactly  $k$  times on  $[0, L]$  then*

- (1)  $\hat{X}_a$  explodes exactly  $k$  times to  $+\infty$  on  $[0, L]$ ,
- (2) for every  $i \in \{1, \dots, k\}$ , we have  $\zeta_a(i - 1) \leq L - \hat{\zeta}_a(k - i + 1) \leq \zeta_a(i)$ .

**Proof** (1) is immediate as the numbers of explosions of  $X_a$  and  $\hat{X}_a$  are related to the number of eigenvalues below  $-a$  of  $\mathcal{H}_L$  and  $\hat{\mathcal{H}}_L$ , and these two operators share the same eigenvalues. To prove the second property, note that almost surely, for all  $t_0 \in \mathbf{Q} \cap (0, L)$  the operators

$$\mathcal{H}_{0,t_0} := -\partial_x^2 + \xi, \quad x \in (0, t_0),$$

and

$$\hat{\mathcal{H}}_{L-t_0,L} := -\partial_x^2 + \hat{\xi}, \quad x \in (L - t_0, L),$$

share the same eigenvalues. Assume that for some  $i, L - \hat{\zeta}_a(k - i + 1) > \zeta_a(i)$  and pick a rational number  $t_0$  in between these two values. Since the diffusion  $X_a$  explodes at least  $i$  times on the interval  $[0, t_0]$ , the operator  $\mathcal{H}_{0,t_0}$  has at least  $i$  eigenvalues below  $-a$ . On the other hand, the diffusion  $\hat{X}_a$  explodes at most  $i - 1$  times on  $[L - t_0, L]$  so that the diffusion  $\hat{X}_a^{L-t_0}$  that starts at  $-\infty$  at time  $L - t_0$  cannot explode more than  $i - 1$  times either. Consequently,  $\hat{\mathcal{H}}_{L-t_0,L}$  has at most  $i - 1$  eigenvalues below  $-a$ . This raises a contradiction. The inequality  $\zeta_a(i - 1) \leq L - \hat{\zeta}_a(k - i + 1)$  follows by symmetry, thus concluding the proof.  $\square$

### 3.2 Fine estimates on the exceptional excursions

In this subsection, we collect some precise estimates on the behavior of the diffusion  $X_a$  during its first exceptional excursion to  $-\sqrt{a}$ .

We let  $\theta_a := \tau_{-\sqrt{a}}(X_a)$  be the first hitting time of  $-\sqrt{a}$ ,  $\nu_a$  be the last hitting time of 0 before  $\theta_a$  and  $\iota_a$  the last hitting time of  $\sqrt{a}$  before  $\theta_a$ . We take similar definitions for the time-reversed diffusion  $\hat{X}_a$ :  $\hat{\theta}_a$  is the first hitting time of  $+\sqrt{a}$ ,  $\hat{\nu}_a$  is the last hitting time of 0 before  $\hat{\theta}_a$  and  $\hat{\iota}_a$  is the last hitting time of  $-\sqrt{a}$  before  $\hat{\theta}_a$ . Note that  $\theta_a < \infty$  a.s. since  $X_a$  explodes in finite time a.s.

We call *excursion* of  $X_a$  a portion of the trajectory that starts at  $\sqrt{a}$ , reaches  $-\sqrt{a}$  while staying (strictly) below  $\sqrt{a}$  and then comes back to  $\sqrt{a}$  (either after an explosion and a restart from  $+\infty$  or without explosion). Note that the diffusion  $X_a$  starts its first excursion at time  $\iota_a$ .

In the next proposition, we control the behavior of  $X_a$  during its first excursion. We refer to Fig. 8 for an illustration of the path  $X_a$ . Recall that  $t_a = \ln a / \sqrt{a}$  and  $h_a = \ln a / a^{1/4}$ .

**Proposition 3.3** (Typical diffusion on its first excursion) *There exists some constant  $C > 0$  such that for all  $a$  large enough, with a probability at least  $1 - O(1/\ln \ln a)$ , the following holds:*

- (i) Oscillations around  $\sqrt{a}$ :

$$\text{For all } t \in [\frac{3}{8}t_a, \tau_{-2\sqrt{a}}(\hat{X}_a)), \quad \int_{\frac{3}{8}t_a}^t X_a(s)ds \in [\sqrt{a} - h_a, \sqrt{a} + h_a].$$

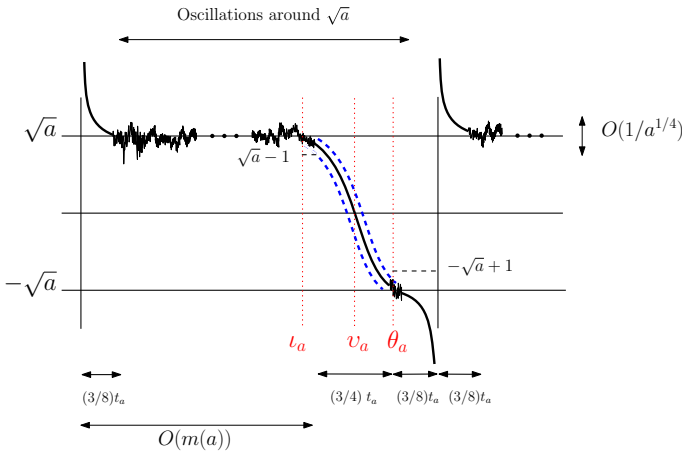


Fig. 8 Schematic path of  $X_a$  and its first excursion to  $-\sqrt{a}$  in the case of an explosion

(ii) Crossing: For all  $t \in [t_a, \theta_a]$ ,

$$|X_a(t) - \sqrt{a} \tanh(-\sqrt{a}(t - v_a))| \leq C \frac{\sqrt{a}}{\ln a},$$

and

$$v_a - t_a \geq (3/8)t_a - C \frac{\ln \ln a}{\sqrt{a}},$$

$$|\theta_a - v_a - (3/8)t_a| \leq C \frac{\ln \ln a}{\sqrt{a}}.$$

(iii) Explosion and/or return to  $\sqrt{a}$ : If after time  $\theta_a$  the diffusion  $X_a$  explodes before coming back to  $+\sqrt{a}$ , then  $X_a$  stays below  $-\sqrt{a} + 1$  in  $[\theta_a, \zeta_a(1))$  and explodes within a time  $(3/8)t_a + (\ln \ln a)^2 / (4\sqrt{a})$ .  
 If it does not explode,  $X_a$  returns to  $\sqrt{a}$  within a time  $(3/4)t_a + C(\ln \ln a)^2 / \sqrt{a}$ .

**Remark 3.4**

- In this proposition we do not bound from above  $v_a - t_a$ , as we do not need to control this time for our purposes.
- In (iii), the time necessary to explode (resp. return to  $\sqrt{a}$ ) is in fact  $(3/8)t_a + O(\ln \ln a / \sqrt{a})$  (resp.  $(3/4)t_a + O(\ln \ln a / \sqrt{a})$ ). The choice of  $(\ln \ln a)^2 / \sqrt{a}$  is arbitrary and any time of the form  $C_a \ln \ln a / \sqrt{a}$  with  $C_a \rightarrow \infty$  would give a probability going to 1.

The next proposition controls the difference between two diffusions  $X_a$  and  $X_{a+\varepsilon}$ , for  $\varepsilon$  not too large, during the first excursion of  $X_a$ . Again, the bounds are not optimal but they suffice for our purposes.

**Proposition 3.5** (Coupling during the excursion) *Let  $\varepsilon = \varepsilon_a \in (a^{-2/3}, 1)$ . There exists some constant  $C > 0$  such that with a probability at least  $1 - O(1/\ln a)$ , the following holds for all a large enough:*

- For all  $t \in [v_a - \frac{1}{16}t_a, v_a + \frac{1}{16}t_a]$ ,  $|X_{a+\varepsilon}(t) - X_a(t)| \leq 1$ ,*
- For all  $t \in [v_a + \frac{1}{16}t_a, \theta_a - \frac{1}{16}t_a]$ ,  $X_{a+\varepsilon}(t) \leq -\sqrt{a} + C a^{3/7}$ ,*
- For all  $t \in [v_a + \frac{1}{16}t_a, \theta_a]$ ,  $X_{a+\varepsilon}(t) \leq \sqrt{a} - 1$ .*

As we will see later on, the above estimates provide a good control on the first eigenfunction over the time-interval  $[t_a, \theta_a]$  (for some well chosen  $a$ ). We will control the first eigenfunction after time  $\theta_a$  by using the time-reversed diffusions. The idea is that after the stopping time  $\theta_a$ , the Brownian motion makes no exceptional event with large probability so that the time reversal  $\hat{X}_a$  should oscillate around  $-\sqrt{a}$ .

**Proposition 3.6** (Control after time  $\theta_a$ ) *For any  $c \geq 1$  and any  $L = L_a \leq c m(a)$  the following holds true with probability greater than  $1 - (\ln a)^{-1/2}$  for all a large enough. If  $L \geq \theta_a + 10 t_a$  and  $\hat{X}_a$  does not explode before time  $L - \theta_a - 10 t_a$ , then it does not explode in the time-interval  $[0, L - \theta_a]$  and we have the bounds:*

$$\hat{X}_a(L - \theta_a) \leq -\sqrt{a} + h_a, \quad \sup_{t \in [\theta_a, L]} \int_{\theta_a}^t \hat{X}_a(L - s) ds \leq -\sqrt{a} + h_a.$$

**Remark 3.7** For any  $\varepsilon = \varepsilon_a \in (a^{-2/3}, 1)$ , a straightforward adaptation of the proof of Proposition 3.6 shows that under the same conditions we have the further bounds:

$$-\sqrt{a} - h_a \leq \hat{X}_{a+\varepsilon}(L - \theta_a), \quad -\sqrt{a} - h_a \leq \inf_{t \in [\theta_a, L - (3/8)t_a]} \int_{\theta_a}^t \hat{X}_{a+\varepsilon}(L - s) ds.$$

with probability greater than  $1 - (\ln a)^{-1/2}$ .

For the sake of readability, we postpone the proofs of those technical propositions to Sect. 5 and proceed to the proof of the main results of this paper.

### 3.3 Preliminaries for the proof of Theorems 1, 2 and 3

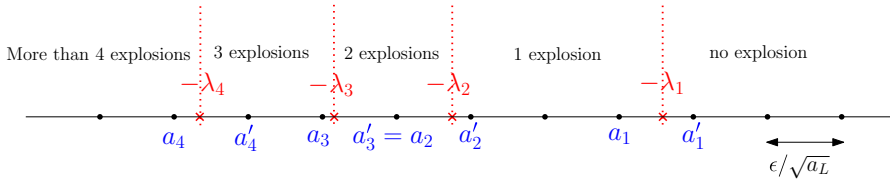
The collection (indexed by  $L$ ) of random variables

$$(4\sqrt{a_L}(\lambda_k + a_L), U_k/L, m_k)_{k \geq 1},$$

is tight since  $\mathcal{Q}_L$  is tight, see Sect. 2, and since the remaining coordinates belong to the product space  $[0, 1]^{\mathbb{N}} \times \mathcal{P}^{\mathbb{N}}$  which is compact. Let  $(\lambda_k^\infty, U_k^\infty, m_k^\infty)_{k \geq 1}$  be the limit of a converging subsequence. We already know that  $(\lambda_k^\infty)_{k \geq 1}$  is a Poisson point process of intensity  $e^x dx$ . Our goal is to show that for any  $k \geq 1$ ,  $(U_1^\infty, \dots, U_k^\infty)$  is i.i.d. uniform over  $[0, 1]$  and independent of  $\mathcal{Q}_\infty$ , that  $m_i^\infty = \delta_{U_i^\infty}$  for every  $i \in \{1, \dots, k\}$  and to obtain the precise description of the eigenfunctions of Theorems 2 and 3.

From now on,  $k$  is fixed. As in Sect. 2.3, we subdivide  $[0, L]$  into  $2^n$  macroscopic sub-intervals  $[t_j^n, t_{j+1}^n]$  where  $t_j^n := j2^{-n}L$  and denote by  $X_a^j := X_a^{t_j^n}$  the diffusion





**Fig. 9** Locations of the (opposite of the) eigenvalues together with their approximations  $a_k$  and  $a'_k$ . Black dots correspond to points of  $M_L$

starting from  $+\infty$  at time  $t_j^n$  and by  $\hat{X}_a^{j+1} := \hat{X}_a^{L-t_{j+1}^n}$  the time-reversed diffusion starting from  $-\infty$  at time  $L - t_{j+1}^n$ . The underlying idea of the proof is that, on every time-interval  $[t_j^n, t_{j+1}^n]$ , the diffusion  $X_a^j$  is a faithful approximation of  $X_a$  until their first explosions and the content of Propositions 3.3, 3.5 and 3.6 can be applied to  $X_a^j$  and  $\hat{X}_a^{j+1}$  (on an interval of length  $2^{-n}L$  instead of  $L$ ).

**Remark 3.8** The aforementioned approximation is supported by the following heuristics: the restriction of every eigenfunction of  $\mathcal{H}_L$  to a small interval surrounding its maximum coincides, up to small errors, to the principal eigenfunction of the operator  $\mathcal{H}_L$  restricted to that interval.

We gather in the following event the ingredients we need to approximate accurately the Riccati transforms of the eigenfunctions by “typical” diffusions. Since we want very precise control simultaneously on several diffusions and since one has to rule out many undesired boundary effects, its definition is quite long. At first reading, one may skip the details and proceed to the next subsection.

For any  $\varepsilon > 0$ , we consider the following discretization of mesh  $\varepsilon/\sqrt{a_L}$  of the interval  $[a_L - 1/(\varepsilon\sqrt{a_L}), a_L + 1/(\varepsilon\sqrt{a_L})]$ :

$$M_L := \left\{ a \in \left[ a_L - \frac{1}{\varepsilon\sqrt{a_L}}, a_L + \frac{1}{\varepsilon\sqrt{a_L}} \right] : \exists n \in \mathbb{Z}, a = a_L + \frac{n\varepsilon}{\sqrt{a_L}} \right\}.$$

From now on, we will use the notation  $t_L := t_{a_L} = \ln(a_L)/\sqrt{a_L}$ . Notice that, for all  $a \in M_L$ , the difference  $t_L - t_a = O(\ln(a_L)/(\varepsilon a_L^2))$  is negligible compared to  $1/\sqrt{a_L}$  so that the estimates already obtained carry through without modifications upon replacing  $t_a$  by  $t_L$ .

Let  $\mathcal{E}(n, \varepsilon)$  be the event on which there exists a random subset

$$\mathcal{A} := \{a_k < a'_k \leq a_{k-1} < a'_{k-1} \leq \dots \leq a_1 < a'_1\}$$

of the grid  $M_L$  such that the following conditions (a), (b) and (c) are fulfilled.

(a) **Squeezing of the  $k$  first eigenvalues:** We have (see Fig. 9)

$$a_k \leq -\lambda_k < a'_k \leq a_{k-1} \leq \dots < a'_2 \leq a_1 \leq -\lambda_1 < a'_1.$$

(b) **Control of the excursions and explosions** For all  $a \in \mathcal{A}$  and all  $j \in \{0, \dots, 2^n - 1\}$  we have:

- (b)–(i) None of the explosion times of  $X_a$  fall at distance less than  $2^{-2n}L$  from any  $t_j^n$ .
- (b)–(ii) Over the time-interval  $[t_j^n, t_{j+1}^n]$ , the diffusions  $X_a$  and  $X_a^j$  start at most one excursion to  $-\sqrt{a}$  that hits  $-\sqrt{a}$  before time  $t_{j+1}^n$ . Over the first and last time-intervals (i.e.  $j = 0$  and  $j = 2^n - 1$ ), the diffusion  $X_a^j$  does not hit  $-\sqrt{a}$ .
- (b)–(iii) If  $X_a^j$  explodes on  $[t_j^n, t_{j+1}^n]$ , then neither  $X_a^{j-1}$  nor  $X_a^{j+1}$  explodes on  $[t_{j-1}^n, t_j^n]$  and  $[t_{j+1}^n, t_{j+2}^n]$ .
- (b)–(iv) If  $X_a^j$  explodes on  $[t_j^n, t_{j+1}^n]$ , then so does  $X_{a_k}^j$  and their explosion times lie at a distance at most  $(\ln \ln a_L)^2 / (3\sqrt{a_L})$  from each other.

and similarly for the time reversed diffusions (up to obvious modifications in the statement due to the identity in law between  $X_a$  and  $-\hat{X}_a$ ).

(c) **Synchronization and typical diffusions** For all  $a \in \mathcal{A}$  and all  $j \in \{0, \dots, 2^n - 1\}$  we have:

- (c)–(i) *Synchronization of  $X_a$  and  $X_a^j$* : if  $t_j^n \in [\zeta_a(l - 1), \zeta_a(l)]$  for some  $l \in \{1, \dots, k\}$  then

$$|X_a(t) - X_a^j(t)| \leq h_a, \quad \forall t \in [t_j^n + \frac{3}{8}t_L, \tau_{-2\sqrt{a}}^l(X_a)],$$

$$|\tau_{-\infty}(X_a^j) - \zeta_a(l)| < \frac{C}{\sqrt{a}},$$

where  $\tau_{-2\sqrt{a}}^l(X_a) := \inf\{t \geq \zeta_a(l - 1) : X_a = -2\sqrt{a}\}$ .

- (c)–(ii) *Typical behavior of  $X_a^j$  and  $X_a$* : The diffusions  $X_a^j$  and  $X_a$  follow the behavior described in Proposition 3.3 and satisfy the conditions of the event  $\mathcal{D}_a^k$ .
- (c)–(iii) *Coupling*: for all  $a' \in \mathcal{A}$ ,  $X_{a'}^j$  and  $X_{a'}^j$  follow the behavior described in Proposition 3.5.
- (c)–(iv) *Time-reversal*: the diffusion  $\hat{X}_a^{j+1}$  follows the behavior described in Proposition 3.6.

and similarly for the time-reversed diffusions.

**Proposition 3.9** *There exists a constant  $C > 0$  such that for all  $\varepsilon$ , we have*

$$\lim_{n \rightarrow \infty} \lim_{L \rightarrow \infty} \mathbb{P}(\mathcal{E}(n, \varepsilon)) \geq 1 - C\varepsilon.$$

The proof of this proposition is postponed to Sect. 5.

**Remark 3.10** It is not difficult to prove that properties (a), (b)–(i), (b)–(ii) and (b)–(iii) hold with large probability thanks to the convergence of the associated point processes. However, the proof of (b)–(iv) is more involved and rely on coupling arguments for different parameters  $a$ 's. Note that (b)–(iii) and (b)–(iv) are used only for the localization of the eigenfunctions  $\varphi_i$  for  $i \geq 2$  proved in paragraph 3.5.

**Remark 3.11** Note that (b)–(ii) implies that  $X_a$  explodes at most once in each interval  $[t_j^n, t_{j+1}^n]$ . Therefore, (b)–(i) together with the synchronization 3.3 imply that  $X_a$  explodes in  $[t_j^n, t_{j+1}^n]$  iff  $X_a^j$  explodes in  $[t_j^n, t_{j+1}^n]$ . We deduce that the total number of explosions of  $X_a$  on  $[0, L]$  equals the sum over  $j$  of the number of explosions of  $X_a^j$  on  $[t_j^n, t_{j+1}^n]$  on the event  $\mathcal{E}(n, \varepsilon)$ .

**Remark 3.12** At a technical level, the symmetries that we rely on are similar in spirit to those appearing in [31] where the invariant measure of the stochastic Allen–Cahn equation is studied.

In the next subsection, we focus on the first eigenfunction: we show that  $m_1^\infty = \delta_{U_1^\infty}$ , that the shape of  $\varphi_1$  around  $U_1$  is asymptotically given by the inverse of a hyperbolic cosine and that  $\varphi_1$  decays at the exponential rate  $\sqrt{a_L}$  from its localization center. In the subsequent subsection, we prove the same results for the  $i$ -th eigenfunction with  $i \in \{2, \dots, k\}$ , together with the behavior of its local maxima and zeros. In the last subsection, we prove that the r.v.  $(U_1^\infty, \dots, U_k^\infty)$  are i.i.d. uniform over  $[0, 1]$  and independent of  $\mathcal{Q}_\infty$ . These three subsections therefore conclude the proof of Theorems 1, 2 and 3.

### 3.4 The first eigenfunction $\varphi_1$

Recall that we denote the Riccati transform of  $\varphi_1$  by  $\chi_1 = X_{-\lambda_1}$ . Let  $a := a_1$  and  $a' := a'_1$  be the approximations of  $-\lambda_1$ . We work deterministically on the event  $\mathcal{E}(n, \varepsilon)$  for some arbitrary  $\varepsilon > 0$ . Recall that  $X_a$  and  $\hat{X}_a$  explode exactly one time on  $[0, L]$  while  $X_{a'}$  and  $\hat{X}_{a'}$  do not explode. Moreover, these diffusions are “typical” in the sense of (b) and (c) of  $\mathcal{E}(n, \varepsilon)$ .

There is an interval  $[t_j^n, t_{j+1}^n]$  that contains the (unique) explosion time of  $X_a$  which occurs in  $[t_j^n + 2^{-2n}L, t_{j+1}^n - 2^{-2n}L]$  by (b)–(i). Using 3.3, we know that  $X_a^j$  and  $X_a$  synchronize after time  $t_j^n + (3/8)t_L$  and that  $X_a^j$  explodes as well (only one time, by (b)–(ii)). Moreover, using Lemma 3.2 on the interval  $[t_j^n, t_{j+1}^n]$ , we deduce that  $\hat{X}_a^{j+1}(L - \cdot)$  explodes one time in  $[t_j^n, t_{j+1}^n]$  and by (b) and (c) again, that the time-interval  $[t_j^n + 2^{-2n}L, t_{j+1}^n - 2^{-2n}L]$  contains the explosion time of  $\hat{X}_a(L - \cdot)$  as well.

The diffusion  $X_a^j$  makes only one excursion to  $-\sqrt{a}$  on the time interval  $[t_j^n, t_{j+1}^n]$  ((b)–(ii)) which therefore explodes to  $-\infty$  before coming back to  $\sqrt{a}$ : let us denote by  $\theta$  its first hitting time of  $-\sqrt{a}$  after time  $t_j^n$ , by  $\nu$  its last hitting time of 0 before time  $\theta$  and by  $\iota$  its last hitting time of  $\sqrt{a}$ . The same holds for  $\hat{X}_a^{j+1}$  and we use the notations  $\hat{\theta}$ ,  $\hat{\nu}$  and  $\hat{\iota}$ . See Fig. 10 for an illustration of the above diffusions.

For technical reasons, we are also led to set

$$\iota_+ := \max(\iota, t_j^n + (3/8)t_L), \quad \hat{\iota}_+ := \max(\hat{\iota}, L - t_{j+1}^n + (3/8)t_L).$$

Indeed, we will apply the synchronization estimates 3.3 which hold true only after time  $t_j^n + (3/8)t_L$ .

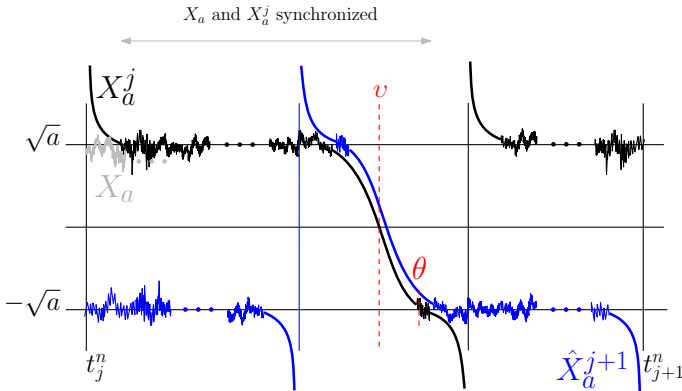


Fig. 10 Typical behavior of  $X_a$ ,  $X_a^j$  and  $\hat{X}_a^{j+1}$  on the time-interval  $[t_j^n, t_{j+1}^n]$

The next proposition gives a lower bound on the first eigenfunction  $\varphi_1$  around  $v$  and the exponential decay after time  $\theta$ .

**Proposition 3.13** *On the event  $\mathcal{E}(n, \varepsilon)$ , the first eigenfunction satisfies the following:*

(1) Lower bound around  $v$ :

$$\forall t \in [t_+, \theta], \quad \frac{\varphi_1(t)}{\varphi_1(v)} \geq \frac{1}{\cosh(\sqrt{a_L}(t-v))} \left(1 - 2C |t-v| \frac{\sqrt{a_L}}{\ln a_L}\right).$$

(2) Exponential decay after time  $\theta$ :

$$\begin{aligned} \forall t \in [\theta, L], \quad & \frac{\varphi_1(t)}{\varphi_1(\theta)} \leq \exp(-(\sqrt{a_L} - \kappa_L)(t - \theta)), \\ \forall t \in \left[\theta, L - \frac{3}{8}t_L\right], \quad & \frac{\varphi_1(t)}{\varphi_1(\theta)} \geq \exp(-(\sqrt{a_L} + \kappa_L)(t - \theta)), \end{aligned}$$

where  $\kappa_L := \ln^2 a_L / a_L^{1/4}$ .

Note that thanks to the time reversal, we get the upper bound:

$$\forall t \in [L - \hat{\theta}, L - \hat{t}_+], \quad \frac{\varphi_1(t)}{\varphi_1(L - \hat{v})} \leq \frac{1}{\cosh(\sqrt{a_L}(t - (L - \hat{v})))} \left(1 + 2C |t - (L - \hat{v})| \frac{\sqrt{a_L}}{\ln a_L}\right), \tag{16}$$

as well as the exponential growth before time  $L - \hat{\theta}$ .

**Proof** For (1), thanks to the synchronization, the Riccati transform  $\chi_1$  is bounded from below by  $X_a^j - h_a$  on the time interval  $[t_+, \theta]$ . The behavior of the first excursion of  $X_a^j$  described in (c)-(ii) yields:

$$\begin{aligned} \frac{\varphi_1(t)}{\varphi_1(v)} &\geq \exp\left(\int_v^t X_a^j(s)ds - h_a|t - v|\right) \\ &\geq \exp\left(\int_v^t \sqrt{a} \tanh(-\sqrt{a}(s - v))ds - \left(C \frac{\sqrt{a_L}}{\ln a_L} + h_a\right)|t - v|\right) \\ &\geq \frac{1}{\cosh(\sqrt{a_L}(t - v))} \left(1 - 2C|t - v| \frac{\sqrt{a_L}}{\ln a_L}\right), \end{aligned}$$

which gives the inequality (1).

For (2): Let us first prove the exponential decay until  $L - (3/8)t_L$ . By (c)-(ii), the only explosion of  $X_a^j$  on  $[t_j^n, t_{j+1}^n]$  occurs at time  $\theta + (3/8)t_L(1 + o(1))$ . Applying Lemma 3.2, we deduce that  $\hat{X}_a^{j+1}(L - \cdot)$  does not explode on  $[\theta + 10t_L, t_{j+1}^n]$  so that (c)-(iv) yields the bound

$$\int_{\theta}^t \hat{X}_a^{j+1}(L - s)ds \leq -\sqrt{a} + h_a, \quad t \in [\theta, t_{j+1}^n - (3/8)t_L]. \tag{17}$$

We thus get the following bound for all  $t \in [\theta, L - (3/8)t_L]$  and all  $L$  large enough:

$$\int_{\theta}^t \hat{X}_a(L - s)ds \leq -(\sqrt{a_L} - \kappa_L/2)(t - \theta).$$

Indeed, for  $t \leq t_{j+1}^n - \frac{3}{8}t_L$  this is a direct consequence of (17) and of the synchronization estimate. For  $t > t_{j+1}^n - \frac{3}{8}t_L$ , it comes from (17), the synchronization estimate, the typical behavior of  $\hat{X}_a$ , the fact that  $t_{j+1}^n - \theta > 2^{-2n}L$  and the simple calculation:

$$\begin{aligned} \int_{\theta}^t \hat{X}_a(L - s)ds &= \int_{\theta}^{t_{j+1}^n - \frac{3}{8}t_L} \hat{X}_a(L - s)ds + \int_{t_{j+1}^n - \frac{3}{8}t_L}^{L - \frac{3}{8}t_L} \hat{X}_a(L - s)ds - \int_t^{L - \frac{3}{8}t_L} \hat{X}_a(L - s)ds \\ &\leq -(\sqrt{a_L} - 2h_{a_L})\left(t_{j+1}^n - \frac{3}{8}t_L - \theta\right) - (\sqrt{a_L} - h_{a_L})\left(L - t_{j+1}^n\right) \\ &\quad + (\sqrt{a_L} + h_{a_L})\left(L - \frac{3}{8}t_L - t\right) \\ &\leq -(\sqrt{a_L} - \kappa_L/2)(t - \theta), \end{aligned} \tag{18}$$

for all  $L$  large enough. Since  $\chi_1(t) \leq \hat{X}_a(L - t)$  for all  $t \in [\theta, L)$ , we deduce the upper bound of the first inequality until time  $L - (3/8)t_L$  but with the improved rate  $\sqrt{a_L} - \kappa_L/2$ .

For the lower bound, the proof is similar except that one has to replace  $\hat{X}_a$  and  $\hat{X}_a^j$  by  $\hat{X}_{a'}$  and  $\hat{X}_{a'}^j$  and use Remark 3.7.

For the interval  $[L - (3/8)t_L, L]$ , thanks to (c)-(ii) and using the fact that  $\hat{X}_{a'}(t) \leq \chi_1(L - t) \leq \hat{X}_a(t)$  for all  $t \in (0, \tau_{+\infty}(\hat{X}_a))$  we get

$$\sup_{t \in (0, (3/8)t_L]} \left| \chi_1(L - t) + \sqrt{a_L} \coth(\sqrt{a_L}(L - t)) \right| \leq 1.$$

We deduce that

$$\forall t \in [L - (3/8)t_L, L], \quad \varphi_1(t) \leq \varphi_1\left(L - \frac{3}{8}t_L\right) \exp\left(-(\sqrt{a_L} - 1)\left(t - L + \frac{3}{8}t_L\right)\right),$$

and the desired upper bound follows, using the exponential decay until  $L - (3/8)t_L$  already established but with an improved rate, and  $L - \theta \geq 2^{-2n}L - O(t_L)$  (thanks to (b)–(i) and (c)–(ii)). □

We will now show that the times  $\nu$ ,  $L - \hat{\nu}$  and  $U_1$  are very close to each other so that the time intervals  $[\nu, \theta]$  and  $[L - \hat{\nu}, L - \hat{\nu}_+]$  overlap and the previous proposition thus describes the behavior of the first eigenfunction around its maximum. This is the key point and most difficult part of the proof. It relies on the ordering of the forward/backward diffusions, the coupling of  $X_a$  and  $X_{a'}$  and the exponential decays of Proposition 3.13.

**Proposition 3.14** *On the event  $\mathcal{E}(n, \varepsilon)$ , the eigenfunction  $|\varphi_1|$  reaches its maximum at distance at most  $2C/(\sqrt{a_L} \ln a_L)$  from  $\nu$  and  $L - \hat{\nu}$ . Moreover, for all  $t \in [L - \hat{\theta}, \theta]$ , we have*

$$\frac{\varphi_1(t)}{\varphi_1(U_1)} = \frac{1}{\cosh(\sqrt{a_L}(t - U_1))} \left(1 + O\left(|t - U_1| \frac{\sqrt{a_L}}{\ln a_L}\right) + O\left(\frac{1}{\ln a_L}\right)\right) \quad (19)$$

**Proof** Let us first prove that all the points where the eigenfunction  $|\varphi_1|$  reaches its maximum over the time-interval  $[\nu - (1/16)t_L, L]$  lie at distance at most  $2C/(\sqrt{a_L} \ln a_L)$  from  $\nu$ . Indeed:

- Using (c), for all  $t \in [\nu - (1/16)t_L, \nu + (1/16)t_L]$ , we have  $X_a^j - 1 \leq \chi_1(t) \leq X_{a'}^j(t) \leq X_a^j + 1$  and:

$$-C \frac{\sqrt{a_L}}{\ln a_L} - 1 \leq \chi_1(t) - \sqrt{a_L} \tanh(-\sqrt{a_L}(t - \nu)) \leq C \frac{\sqrt{a_L}}{\ln a_L} + 1. \quad (20)$$

This implies that all the points where  $|\varphi_1|$  reaches its maximum over  $[\nu - (1/16)t_L, \nu + (1/16)t_L]$  lie at distance at most  $2C/(\sqrt{a_L} \ln a_L)$  from  $\nu$ .

- By (c)–(iii), the diffusion  $X_{a'}^j$  is below  $-\sqrt{a}(1 + o(1))$  on the interval  $[\nu + (1/16)t_L, \theta - (1/16)t_L]$ . Therefore,  $|\varphi_1|$  decreases on this time interval and we have the bound

$$|\varphi_1\left(\theta - \frac{1}{16}t_L\right)| \leq |\varphi_1\left(\nu + \frac{1}{16}t_L\right)| \exp\left(-\sqrt{a}(1 + o(1))t_L\left(\frac{1}{4} + o(1)\right)\right),$$

where we used  $\theta - \nu = (3/8)t_L(1 + o(1))$ . Since  $X_{a'}^j$  stays below  $\sqrt{a}$  on  $[\theta - (1/16)t_L, \theta]$ , we deduce that  $|\varphi_1|$  remains below  $|\varphi_1(\nu + \frac{1}{16}t_L)|$  on the time interval  $[\nu + (1/16)t_L, \theta]$ .

- From Proposition 3.13, the eigenfunction  $|\varphi_1|$  decays exponentially on the time-interval  $[\theta, L]$ .

By time-reversal, the eigenfunction  $|\varphi_1|$  reaches its maximum over the time-interval  $[0, L - \hat{v} + (1/16)t_L]$  at distance at most  $2C/(\sqrt{a_L} \ln a_L)$  from  $L - \hat{v}$ .

Note that in the time-interval  $[L - \hat{\tau}_{+\infty}(\hat{X}_a^{j+1}), \tau_{-\infty}(X_a^j)]$ , the diffusion  $\hat{X}_a^{j+1}(L - \cdot)$  is bounded from below by  $X_a^j(\cdot)$  as the Riccati transform of the first eigenfunction of the operator

$$-\partial_x^2 + \xi, \quad x \in (t_j^n, t_{j+1}^n),$$

stays in between those two diffusions (see Fig. 10). Using that the diffusion  $X_a^j$  is bounded from below by  $t \mapsto \sqrt{a} \tanh(-\sqrt{a}(t - v)) - C\sqrt{a}/\ln a$  on the time interval  $[t, v]$ , we deduce that  $\hat{X}_a^{j+1}(L - \cdot)$  cannot vanish on the time-interval  $[t, v - 2C/(\sqrt{a} \ln a)]$ . As it reaches  $\sqrt{a}$  in  $[t, t_{j+1}^n]$  and makes at most one excursion to  $\sqrt{a}$  on  $[t_j^n, t_{j+1}^n]$ , we deduce that  $L - \hat{\theta}$  is larger than  $v$  and  $L - \hat{v}$  is larger than  $v - 2C/(\sqrt{a_L} \ln a_L)$ . This implies that the maximum over the time-interval  $[v - (1/16)t_L, L]$  is a maximum over the whole interval  $[0, L]$ . Therefore  $U_1$  and  $v$  are at distance smaller than  $2C/(\sqrt{a_L} \ln a_L)$  from each other and the same holds for  $L - \hat{v}$ .

Using (20), we obtain for all  $t \in [v - (1/16)t_L, v + (1/16)t_L]$

$$\frac{1}{\cosh(\sqrt{a_L}(t-v))} \left(1 - 2C |t-v| \frac{\sqrt{a_L}}{\ln a_L}\right) \leq \frac{\varphi_1(t)}{\varphi_1(v)} \leq \frac{1}{\cosh(\sqrt{a_L}(t-v))} \left(1 + 2C |t-v| \frac{\sqrt{a_L}}{\ln a_L}\right),$$

and therefore  $\varphi_1(U_1)/\varphi_1(v)$  and  $\varphi_1(U_1)/\varphi_1(L - \hat{v})$  are both of order  $1 + O(1/\ln^2 a_L)$ . Note that  $t_j^n + (3/8)t_L < L - \hat{\theta}$  since otherwise there would be an explosion of  $\hat{X}_a(L - \cdot)$  close to  $t_j^n$  thus violating (b)-(i); similarly we have  $\theta < t_{j+1}^n - (3/8)t_L$ . Then, combining the lower bound of Proposition 3.13-(1) and its time-reversed version (16), together with the inequalities  $\iota_+ < L - \hat{\theta} < \theta < L - \hat{\iota}_+$  we obtain (19).  $\square$

We now conclude the proof of the statements of Theorems 1 and 2 regarding the first eigenfunction (except the statement on the law of  $U_1^\infty$ , which is presented in Sect. 3.6). We work on the event  $\mathcal{E}(n, \varepsilon)$ :

• **Exponential decay**

The simple inequalities  $e^{-|x|} \leq 1/\cosh(x) \leq 2e^{-|x|}$  applied to (19), together with Proposition 3.13-(2) and its time-reversed version give the exponential decay on  $[0, L]$ .

• **Localization around  $U_1$  and shape of the eigenfunction**

Using the exponential decay at rate  $\sqrt{a_L}$ , we easily get

$$m_1([\theta/L, 1]) \leq \varphi_1(\theta)^2 O\left(\frac{1}{\sqrt{a_L}}\right).$$

Moreover, the time reversal gives:

$$m_1([0, (L - \hat{\theta})/L]) \leq \varphi_1(L - \hat{\theta})^2 O\left(\frac{1}{\sqrt{a_L}}\right).$$

Integrating (19) of Proposition 3.14, we obtain

$$m_1([L - \hat{\theta}/L, \theta/L]) = \varphi_1(U_1)^2 \frac{2}{\sqrt{a_L}} \left(1 + O\left(\frac{1}{\ln a_L}\right)\right).$$

Moreover, using that  $\theta - U_1 = (3/8)t_L(1 + o(1))$ , we get

$$\frac{\varphi_1(\theta)}{\varphi_1(U_1)} \leq \exp(-(3/8) \ln a_L(1 + o(1))),$$

and similarly  $|\varphi_1(L - \hat{\theta})| \leq |\varphi_1(U_1)| \exp(-(3/8) \ln a_L(1 + o(1)))$ .

As  $m_1$  is a probability measure, we deduce that  $\varphi_1^2(U_1) = (\sqrt{a_L}/2)(1 + o(1))$  and the statement about the shape of the first eigenfunction follows. To obtain the shape of the Brownian motion, it suffices to combine the identity

$$\chi_1(t) = \chi_1(U_1) + \int_{U_1}^t (-\lambda_1 - \chi_1(s)^2) ds + B(t) - B(U_1),$$

with the estimate (20), noticing that  $|a_L + \lambda_1| = O(1/\sqrt{a_L})$ .

It is then easy to see using again (19) that any interval centered around  $U_1$  and of length much greater than  $1/\sqrt{a_L}$  has a mass going to 1 as  $L \rightarrow \infty$  so that  $m_1$  is asymptotically as close as desired to a Dirac mass at  $U_1/L$ .

### 3.5 The next eigenfunctions

We now turn to the  $i$ -th eigenfunction, for some  $i \in \{2, \dots, k\}$ . Let  $a = a_i$  and  $a' = a'_i$ . Again, we work deterministically on the event  $\mathcal{E}(n, \varepsilon)$ .

The main idea is that the trajectory  $\chi_i$  follows the forward diffusions  $X_{a_i}$  and  $X_{a'_i}$  until the additional explosion of  $X_{a_i}$ . It then follows the time-reversed diffusions  $\hat{X}_{a_i}$  and  $\hat{X}_{a'_i}$  up to time  $L$  (see Fig. 11).

There exist  $j_1 < j_2 < \dots < j_i$  such that the  $i$  explosion times of  $X_a$  lie in the intervals  $[t_{j_\ell}^n, t_{j_{\ell+1}}^n)$ ,  $\ell = 1, \dots, i$ . Each  $X_a^{j_\ell}$  explodes once, by (b)–(i) and 3.3, and makes only one excursion to  $-\sqrt{a}$ , by (b)–(ii). We then call  $\theta_\ell$  the location of its first hitting time of  $-\sqrt{a}$  (resp.  $\hat{\theta}_\ell$  for the time-reversed diffusions). There exists a unique  $i_* \in \{1, \dots, i\}$  such that, for  $j_* = j_{i_*}$ ,  $X_a^{j_*}$  explodes but  $X_{a'}^{j_*}$  does not. Without loss of generality, we can suppose that  $i_* \geq 2$ : indeed, the case  $i_* = 1$  corresponds to the case  $i_* = i$  upon reversing time.

Let us now study  $\varphi_i$ . We denote by  $0 = z_0 < z_1 < \dots < z_i = L$  the zeros of  $\varphi_i$ , or equivalently the explosion times of  $\chi_i$ . We also let  $x_\ell$  be the (first if many) point where  $|\varphi_i|$  reaches its maximum on the interval  $[z_{\ell-1}, z_\ell]$ .



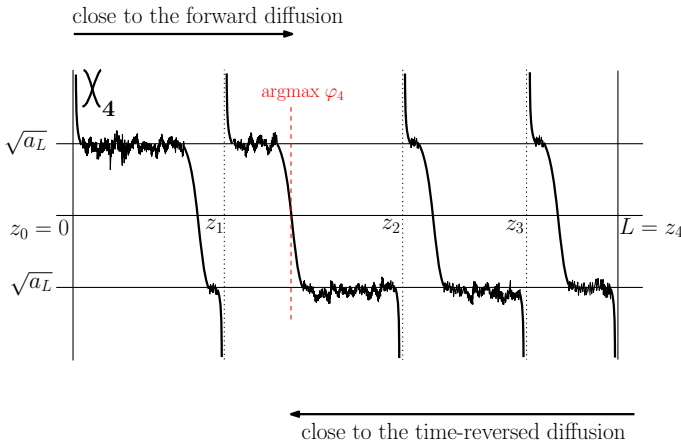


Fig. 11 Schematic path of  $\chi_4$

We first describe the behavior of  $\varphi_i$  on  $[z_{\ell-1}, z_\ell]$  with  $\ell < i_*$ .

**Lemma 3.15** *Let  $\ell < i_*$ . On the event  $\mathcal{E}(n, \varepsilon)$ , we have:*

(I) *Deterministic entrance and explosion:*

$$\forall t \in \left[ z_{\ell-1}, z_{\ell-1} + \frac{3}{8}t_L \right], \quad \varphi_i(t) = \varphi'_i(z_{\ell-1}) \frac{\sinh(\sqrt{a_L}(t - z_{\ell-1}))}{\sqrt{a_L}} (1 + o(1)),$$

$$\forall t \in \left[ z_\ell - \frac{3}{8}t_L, z_\ell \right], \quad \varphi_i(t) = \varphi'_i(z_\ell) \frac{\sinh(\sqrt{a_L}(t - z_\ell))}{\sqrt{a_L}} (1 + o(1)).$$

(II) *Exponential decay: For all  $t \in [z_{\ell-1} + \frac{3}{8}t_L, z_\ell - \frac{3}{8}t_L]$*

$$\frac{1}{100} \exp\left(-\left(\sqrt{a_L} + \frac{1}{2}\kappa_L\right)|t - x_\ell|\right) \leq \frac{\varphi_i(t)}{\varphi_i(x_\ell)} \leq 100 \exp\left(-\left(\sqrt{a_L} - \frac{1}{2}\kappa_L\right)|t - x_\ell|\right),$$

where  $\kappa_L := \ln^2 a_L / a_L^{1/4}$ .

(III) *Inverse of hyperbolic cosine around  $x_\ell$ :*

$$\forall t \in [L - \hat{\theta}_\ell, \theta_\ell], \quad \frac{\varphi_i(t)}{\varphi_i(x_\ell)} = \frac{1}{\cosh(\sqrt{a_L}(t - x_\ell))} \left( 1 + O\left(|t - x_\ell| \frac{\sqrt{a_L}}{\ln a_L}\right) + O\left(\frac{1}{\ln a_L}\right) \right).$$

(IV) *We have the bounds*

$$|x_\ell + 3/8t_L - \theta_\ell| \leq \frac{5C \ln \ln a_L}{\sqrt{a_L}}, \quad |z_\ell - \theta_\ell - 3/8t_L| \leq \frac{2}{3} \frac{(\ln \ln a_L)^2}{\sqrt{a_L}}.$$

**Proof** We treat in detail the case  $\ell = 1$ . The other cases follow from the same arguments, one simply has to notice that the diffusions  $X_a$  and  $X_{a'}$  have a delay at the

starting time of the interval, but since this delay is negligible compared to  $t_L$  the proof carries through (indeed, the only probabilistic estimate in the proof of Lemma 4.2 is a control of the Brownian motion on  $[0, t_L]$ ).

Set  $j = j_1$ . Let  $\theta_a, \theta_{a'}$  be the first hitting times of  $-\sqrt{a}, -\sqrt{a'}$  by  $X_a^j$  and  $X_{a'}^j$ , and similarly for  $\nu_a, (\nu_a)_+, \nu_{a'}, (\nu_{a'})_+$ . Moreover, let  $\hat{\theta}_a, \hat{\theta}_{a'}$  be the first hitting times of  $\sqrt{a}, \sqrt{a'}$  by  $\hat{X}_a^{j+1}$  and  $\hat{X}_{a'}^{j+1}$ . Applying the arguments of the proof of Proposition 3.13-(1) we obtain

$$\forall t \in [(\nu_a)_+, \theta_a], \quad \frac{\varphi_i(t)}{\varphi_i(\nu_a)} \geq \frac{1}{\cosh(\sqrt{a_L}(t - \nu_a))} \left(1 - 2C |t - \nu_a| \frac{\sqrt{a_L}}{\ln a_L}\right).$$

Applying the arguments of the proof of the *time-reversed* version of Proposition 3.13-(2) we obtain

$$\begin{aligned} \forall t \in [0, L - \hat{\theta}_a], \quad & \frac{\varphi_i(t)}{\varphi_i(L - \hat{\theta}_a)} \leq \exp\left(-\left(\sqrt{a_L} - \frac{1}{2}\kappa_L\right)(L - \hat{\theta}_a - t)\right), \\ \forall t \in [(3/8)t_L, L - \hat{\theta}_a], \quad & \frac{\varphi_i(t)}{\varphi_i(L - \hat{\theta}_a)} \geq \exp\left(-\left(\sqrt{a_L} + \frac{1}{2}\kappa_L\right)(L - \hat{\theta}_a - t)\right). \end{aligned} \tag{21}$$

Note that, compared to the original version of the bound, here we have an additional factor  $1/2$  in front of  $\kappa_L$  but in the proof, this prefactor can actually be chosen arbitrarily. This is enough to deduce that the maximum of  $|\varphi_i|$  over  $[0, L - \hat{\theta}_a]$  is achieved at  $L - \hat{\theta}_a$ .

The proof now deviates from that presented in the previous subsection for the first eigenfunction. Indeed, we are in a situation where  $X_a$  and  $X_{a'}$  remain close to each other and explode roughly at the same time (while the latter did not explode in the previous case). More precisely, by (b)–(iv) the explosions times of  $X_a^j$  and  $X_{a'}^j$  lie at distance at most  $(\ln \ln a_L)^2 / (3\sqrt{a_L})$  from each other. Since with (c)–(ii), their behaviors are almost deterministic on a time-interval of length  $(3/8 + 3/4)t_L$  before their explosion times and since they are very close to each other on  $[\nu_a - (1/16)t_L, \nu_a + (1/16)t_L]$  we deduce that

$$0 \leq \nu_{a'} - \nu_a \leq 2C / (\ln a_L \sqrt{a_L}), \quad \text{which implies} \quad 0 \leq \theta_{a'} - \theta_a \leq 3C \ln \ln a_L / \sqrt{a_L}. \tag{22}$$

By monotonicity  $X_a \leq \mathcal{X}_i \leq X_{a'}$  until the first explosion time of  $X_a$ . Using the synchronization estimates 3.3, together with (c)–(ii) on  $X_a^j$  and  $X_{a'}^j$  and (22), we deduce that

$$\forall t \in [(\nu_a)_+, \theta_a], \quad -C \frac{\sqrt{a_L}}{\ln a_L} - 1 \leq \mathcal{X}_i(t) - \sqrt{a_L} \tanh(-\sqrt{a_L}(t - \nu_a)) \tag{23}$$

$$\forall t \in [(\nu_{a'})_+, \theta_{a'}], \quad \mathcal{X}_i(t) - \sqrt{a_L} \tanh(-\sqrt{a_L}(t - \nu_{a'})) \leq 3C \frac{\sqrt{a_L}}{\ln a_L} + 1. \tag{24}$$

In addition,  $X_{a'}^j$  remains below  $-\sqrt{a_L} + 1$  on  $[\theta_{a'}, \zeta_{a'}]$ , and it cannot be higher than  $-\sqrt{a_L} + 2C\sqrt{a_L}/\ln a_L$  on the interval of time  $[\theta_a, \theta_{a'}]$  by (c)–(ii) and (22). We thus deduce that  $\chi_i(t) \leq -\sqrt{a_L} + 2C\sqrt{a_L}/\ln a_L$  for all  $t \in [\theta_a, z_1]$ . Therefore all the points where the maximum of  $|\varphi_i|$  over  $[(\iota_a)_+ \vee (\iota_{a'})_+, z_1]$  is achieved lie at distance at most  $4C/(\ln a_L \sqrt{a_L})$  from  $v_a$ .

Let us now prove that the intervals  $[0, L - \hat{\theta}_a]$  and  $[(\iota_a)_+ \vee (\iota_{a'})_+, z_1]$  overlap (here the argument is the same as in the previous subsection). To that end, we observe that  $X_{a'}^j(t) \leq \hat{X}_a^{j+1}(L - t)$  for all  $t \in [L - \hat{\tau}_{+\infty}(\hat{X}_a^{j+1}), \tau_{-\infty}(X_{a'}^j)]$  and that the latter interval is not empty (these two assertions follow from Lemma 3.2 and from the fact that the Riccati transform of the first eigenfunction of  $-\partial_x^2 + \xi$  restricted to  $(t_j^n, t_{j+1}^n)$  lies in between these two diffusions). Henceforth  $L - \hat{\theta}_a > \iota_{a'}$  and  $L - \hat{v}_a > v_{a'} - C/(\ln a_L \sqrt{a_L}) > v_a - C/(\ln a_L \sqrt{a_L})$ . Moreover,  $L - \hat{\theta}_a > t_j^n + (3/8)t_L$  since otherwise  $\hat{X}_a(L - \cdot)$  would explode close to  $t_j^n$ . Therefore,  $L - \hat{\theta}_a > (\iota_{a'})_+$ . The same argument applies upon replacing  $X_{a'}^j$  by  $X_a^j$  and yields the inequality  $L - \hat{\theta}_a > (\iota_a)_+$  which concludes the proof of the overlapping.

Consequently, all the points where the maximum of  $|\varphi_i|$  over  $[0, z_1]$  is achieved lie at distance at most  $4C/(\ln a_L \sqrt{a_L})$  from  $v_a$ . This proves the first bound of (IV). The second bound of (IV) is a consequence of the squeezing  $X_a \leq \chi_i \leq X_{a'}$ , of 3.3 and (c)–(ii) and of (22).

Assertion (III) is deduced from  $|x_1 - v_a| \leq 4C/(\ln a_L \sqrt{a_L})$ , combined with (23), (24) and the inequality  $(\iota_a)_+ \vee (\iota_{a'})_+ \leq L - \hat{\theta}_a$ . It also ensures that for all  $t \in [L - \hat{\theta}_a, \theta_a]$

$$\frac{1}{10} \exp\left(-\left(\sqrt{a_L} + \frac{1}{2}\kappa_L\right)|t - x_1|\right) \leq \frac{\varphi_i(t)}{\varphi_i(x_1)} \leq 10 \exp\left(-\left(\sqrt{a_L} - \frac{1}{2}\kappa_L\right)|t - x_1|\right).$$

Using (21), these inequalities still hold on  $[(3/8)t_L, \theta_a]$ . It remains to control the decay on  $[\theta_a, z_1 - (3/8)t_L]$ : by the second bound of (IV) this interval has length at most  $(\ln \ln a_L)^2/\sqrt{a_L}$ . Furthermore on this interval of time, we have already seen that  $\chi_i$  remains below  $-\sqrt{a_L} + 2C\sqrt{a_L}/\ln a_L$  and by (c)–(ii) it remains above  $-\sqrt{a_L} - 2$ . Henceforth, for all  $t \in [\theta_a, z_1 - (3/8)t_L]$

$$\frac{1}{2} \exp\left(-\sqrt{a_L}(t - \theta_a)\right) \leq \frac{\varphi_i(t)}{\varphi_i(\theta_a)} \leq 2 \exp\left(-\sqrt{a_L}(t - \theta_a)\right).$$

This is enough to get (II).

Recall that  $X_a(t) \leq \chi_i(t)$  for all  $t \in (0, \zeta_a(1))$  and  $\chi_i(t) \leq X_{a'}(t)$  for all  $t \in (0, z_1)$ . The deterministic behavior of  $X_a$  and  $X_{a'}$  near their starting point gives:

$$\sup_{t \in (0, (3/8)t_L]} \left| \chi_i(t) - \sqrt{a_L} \coth(\sqrt{a_L}t) \right| \leq 1. \tag{25}$$

Since  $\varphi_i'(t) = \chi_i(t)\varphi_i(t)$ , we deduce that for all  $0 < t_0 < t < \tau_{-\infty}(\chi_i)$  we have

$$\varphi_i(t) = \frac{\varphi_i'(t_0)}{\chi_i(t_0)} e^{\int_{t_0}^t \chi_i(s) ds}.$$

Using (25) and passing to the limit as  $t_0 \downarrow 0$ , we deduce that for all  $t \in (0, \frac{3}{8}t_L]$ , we obtain

$$\varphi_i(t) = \varphi'_i(0) \frac{\sinh(\sqrt{a_L}t)}{\sqrt{a_L}}(1 + O(t_L)). \tag{26}$$

Regarding the behavior near  $z_1$ : both  $X_a$  and  $X_{a'}$  explode at a distance negligible compared to  $t_L$  from each other thanks to (b)–(iv). The first explosion time of  $\chi_i$  necessarily lies in between those two explosion times. The proof of Lemma 4.3 then ensures that if we let  $y = -\sqrt{a_L} - h_L$  then

$$\forall t \in [\tau_y(\chi_i), z_1], \quad \left| \chi_i(t) + \sqrt{a_L} \coth(\sqrt{a_L}(z_1 - t)) \right| \leq 1.$$

Indeed, the only probabilistic ingredient in that proof is a control on the Brownian motion on an interval of size  $t_L$  before the explosion time: on the event  $\mathcal{E}(n, \varepsilon)$ , we have this control before the explosion times of  $X_{a'}$  and  $X_a$  so that the control holds true before the explosion time of  $\chi_i$ . A similar calculation as above then shows that for all  $t \in [\tau_y(\chi_i), z_1]$ , we have

$$\varphi_i(t) = \varphi'_i(z_1) \frac{\sinh(\sqrt{a_L}(t - z_1))}{\sqrt{a_L}}(1 + O(t_L)).$$

Since  $z_1 - \tau_y(\chi_i) \geq (3/8)t_L - \ln \ln a_L / \sqrt{a_L}$ , and since  $\chi_i(t) \leq -\sqrt{a_L} + 2C\sqrt{a_L} / \ln a_L$  for all  $t \in [\theta_a - \ln \ln a_L / \sqrt{a_L}, \tau_y(\chi_i)]$  we deduce that for all  $t \in [z_1 - (3/8)t_L, z_1]$

$$\varphi_i(t) = \varphi'_i(z_1) \frac{\sinh(\sqrt{a_L}(t - z_1))}{\sqrt{a_L}}(1 + o(1)).$$

□

Let us now describe what happens on the interval  $[z_{i_*-1}, z_{i_*}]$ .

**Lemma 3.16** *On the event  $\mathcal{E}(n, \varepsilon)$ , the eigenfunction  $|\varphi_i|$  satisfies (I)–(III) of Lemma 3.15 on the time interval  $[z_{i_*-1}, z_{i_*}]$ . Moreover all the points where the maximum over  $[z_{i_*-1}, z_{i_*}]$  is reached lie at distance at most  $4C / (\ln a_L \sqrt{a_L})$  from each other, and we have  $|x_{i_*} + 3/8t_L - \theta_{i_*}| \leq 5C \ln \ln a_L / \sqrt{a_L}$ .*

**Proof** The proof follows from the same arguments as in the case of the first eigenfunction, one simply has to take into account the delay between  $X_a$  and  $X_{a'}$  at the beginning of the interval but this delay is negligible compared to  $t_L$ . □

Applying successively Lemma 3.15 and its time-reversed version, together with Lemma 3.16, a straightforward calculation then shows that for all  $t \in [0, L]$

$$\frac{1}{100} \exp(-(\sqrt{a_L} + \kappa_L)|t - x_{i_*}|) 1_{\{t \in D\}} \leq \frac{\varphi_i(t)}{\varphi_i(x_{i_*})} \leq 100 \exp(-(\sqrt{a_L} - \kappa_L)|t - x_{i_*}|),$$

where the set  $D$  is defined in the statement of Theorem 2. (Note that the improved exponential rate  $\sqrt{a_L} - (1/2)\kappa_L$  allows to compensate the “absence” of decay near the zeros).

Recall that by (b)–(iii),  $|z_\ell - x_{i_*}| > 2^{-n}L$  for all  $\ell \neq i_*$ . Hence the maximum of  $|\varphi_i|$  on  $[0, L]$  is  $|\varphi_i(x_{i_*})|$  and necessarily  $U_i = x_{i_*}$ . It is therefore easy to see that

$$m_i\left([0, 1] \setminus \left[\frac{z_{i_*-1}}{L}, \frac{z_{i_*}}{L}\right]\right) \lesssim \varphi_i(x_{i_*})^2 L e^{-\frac{1}{2}\sqrt{a_L}2^{-n}L}.$$

By Lemma 3.16, we also get

$$m_i\left(\left[\frac{z_{i_*-1}}{L}, \frac{z_{i_*}}{L}\right]\right) = \varphi_i(x_{i_*})^2 \frac{2}{\sqrt{a_L}}(1 + o(1)).$$

Therefore,

$$m_i([0, 1]) = \varphi_i(x_{i_*})^2 \frac{2}{\sqrt{a_L}}(1 + o(1)).$$

Since  $m_i$  is a probability measure, we deduce that

$$|\varphi_i(x_{i_*})| \sim a_L^{1/4}/\sqrt{2}.$$

Finally, we claim that for all  $\ell \in \{1, \dots, i\}$  we have the bound

$$|x_\ell - U_{\sigma(\ell)}| \leq 8C/(\ln a_L \sqrt{a_L}).$$

The cases  $\ell < i_*$  and  $\ell > i_*$  are symmetric, so we only consider the former. Let us provide the arguments in the case  $\ell = 1$ , using the notations of the proof of Lemma 3.15 (in particular  $j = j_1$ ). If we let  $p$  be the smallest integer such that  $X_{a_p}^j$  explodes on  $[t_j^n, t_{j+1}^n]$  but  $X_{a_p}^j$  does not, then  $v_{a_p}$  lies at a distance at most  $2C/(\sqrt{a_L} \ln a_L)$  from  $U_{\sigma(1)}$  and from  $v_a$ . Since  $x_1$  lies at a distance at most  $4C/(\sqrt{a_L} \ln a_L)$  from  $v_a$ , the claim follows. Note that  $U_{\sigma(i_*)} = U_i$ .

The rest of the statements of Theorems 1, 2 and 3 regarding the  $i$ -th eigenfunction directly follows from Lemmas 3.15 and 3.16, except the statement on the law of  $U_i^\infty$  which will be proven in the next subsection.

### 3.6 Convergence towards uniform r.v.

**Proposition 3.17** *The r.v.  $(U_1^\infty, \dots, U_k^\infty)$  are i.i.d. uniform over  $[0, 1]$  and independent of  $\mathcal{Q}_\infty$ .*

This proposition, combined with the results of Sect. 2.3, show that  $(\lambda_i^\infty, U_i^\infty)_{i \geq 1}$  is distributed as a Poisson point process on  $\mathbf{R} \times [0, 1]$  with intensity  $e^x dx \otimes dt$ . Indeed,

since the intensity is a product measure, the mere fact that  $(\lambda_i^\infty)_{i \geq 1}$  and  $(U_i^\infty)_{i \geq 1}$  are independent and have the right distributions yields the desired property.

Before we proceed to the proof, we prove a simple characterization of i.i.d. uniform r.v. Let  $\sigma$  be a bijection from  $[0, 1]$  into itself and assume that there exists  $n_0 \geq 1$  such that the restriction of  $\sigma$  to every interval of the form  $[s_i^{n_0}, s_{i+1}^{n_0})$  is affine with slope 1, where  $s_i^{n_0} = i2^{-n_0}$  for  $i \in \{0, \dots, 2^{n_0} - 1\}$ . In other terms,  $\sigma$  is completely characterized by its values at the points  $s_i^{n_0}$  and in between we have  $\sigma(x) = \sigma(s_i^{n_0}) + (x - s_i^{n_0})$ ,  $x \in [s_i^{n_0}, s_{i+1}^{n_0})$ .

**Lemma 3.18** *Let  $V_1, \dots, V_k$  be  $[0, 1]$ -valued r.v., assumed to be almost surely distinct. Suppose that for all bijections  $\sigma$  as above and for all  $j_1, \dots, j_k \in \{0, \dots, 2^{n_0} - 1\}$  we have*

$$\begin{aligned} \mathbb{P}[V_1 \in [s_{j_1}^{n_0}, s_{j_1+1}^{n_0}), \dots, V_k \in [s_{j_k}^{n_0}, s_{j_k+1}^{n_0})] \\ = \mathbb{P}[\sigma(V_1) \in [s_{j_1}^{n_0}, s_{j_1+1}^{n_0}), \dots, \sigma(V_k) \in [s_{j_k}^{n_0}, s_{j_k+1}^{n_0})]. \end{aligned}$$

Then  $(V_1, \dots, V_k)$  is i.i.d. uniform over  $[0, 1]$ .

**Proof** Let  $\pi$  be the law of  $(V_1, \dots, V_k)$ . Let  $D$  be the set of all points  $v = (v_1, \dots, v_k) \in [0, 1]^k$  that have at least two coordinates equal. By assumption,  $\pi(D) = 0$ . Consider the  $2^{kn_0}$  dyadic hypercubes of the type  $[s_{j_1}^{n_0}, s_{j_1+1}^{n_0}) \times \dots \times [s_{j_k}^{n_0}, s_{j_k+1}^{n_0})$ . Call  $D_{n_0}$  the union of all such hypercubes that intersect  $D$ . We have  $D_{n_0} \downarrow D$  as  $n_0 \rightarrow \infty$  so that  $\pi(D_{n_0}) \rightarrow 0$ . On the other hand, the assumption of the statement ensures that  $\pi$  gives the same measure to every hypercube that does not intersect  $D$ . The number of all such hypercubes is  $2^{n_0}(2^{n_0} - 1) \dots (2^{n_0} - k + 1)$ ; notice that this quantity is equivalent to  $2^{kn_0}$  as  $n_0 \rightarrow \infty$ . Since  $\pi([0, 1]^k \setminus D_{n_0}) \rightarrow 1$ , we deduce that the measure of every hypercube that does not intersect  $D$  is equivalent to  $2^{-kn_0}$  as  $n_0 \rightarrow \infty$ . By a simple approximation argument, it is then easy to deduce that  $\pi(A)$  equals the Lebesgue measure of  $A$  for all set  $A$  which is a product of intervals. As a consequence,  $\pi$  is the Lebesgue measure on  $[0, 1]^k$  and the statement of the lemma follows. □

**Proof of Proposition 3.17** Observe that  $\mathcal{Q}_\infty$  is a Poisson point process with intensity  $e^x dx$  that can be written

$$\mathcal{Q}_\infty = \sum_{i \geq 1} \delta_{\lambda_i^\infty},$$

where  $(\lambda_i^\infty)_{i \geq 1}$  is the limit in distribution (for the product topology) of  $((\lambda_i + a_L)/4\sqrt{a_L})_{i \geq 1}$ . Moreover, the r.v.  $U_1^\infty, \dots, U_k^\infty$  are all distinct a.s. Indeed, on the event  $\mathcal{E}(n, \varepsilon)$  the r.v.  $U_1, \dots, U_k$  all lie at a distance at least  $2^{-n}$  from one another so that the r.v.  $U_1^\infty, \dots, U_k^\infty$  are all distinct with probability at least  $1 - O(\varepsilon)$ ; but since  $\varepsilon$  can be taken as small as desired, the latter property holds almost surely.

It suffices to show that  $(U_1^\infty, \dots, U_k^\infty)$  is i.i.d. uniform over  $[0, 1]$  and independent of  $(\lambda_1^\infty, \dots, \lambda_k^\infty)$ . Indeed, since  $k$  is arbitrary, such a result would ensure

that  $(U_1^\infty, \dots, U_k^\infty)$  is independent of  $(\lambda_1^\infty, \dots, \lambda_k^\infty)$  for any  $K \geq k$ , so that  $(U_1^\infty, \dots, U_k^\infty)$  is independent from  $(\lambda_i^\infty)_{i \geq 1}$  as required. If we show that

$$\mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{U_i^\infty \in [s_{j_i}^{n_0}, s_{j_i+1}^{n_0})} f(\lambda_1^\infty, \dots, \lambda_k^\infty) \right] = \mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{\sigma(U_i^\infty) \in [s_{j_i}^{n_0}, s_{j_i+1}^{n_0})} f(\lambda_1^\infty, \dots, \lambda_k^\infty) \right], \tag{29}$$

holds for all continuous functions  $f$  on  $\mathbf{R}^k$  with compact support, all integers  $j_i \in \{0, \dots, 2^{n_0} - 1\}$  and all bijections  $\sigma$  as above, then Lemma 3.18 ensures that  $(U_1^\infty, \dots, U_k^\infty)$  is i.i.d. uniform on  $[0, 1]$ , and classical arguments yield the independence from  $(\lambda_1^\infty, \dots, \lambda_k^\infty)$ .

To prove (27), we proceed as follows. Fix a bijection  $\sigma$  and an integer  $n_0 \geq 1$  as above. Let  $\tilde{\xi} := \xi \circ \sigma_L^{-1}$  where  $\sigma_L : [0, L] \ni x \mapsto L\sigma(x/L)$ . In other words, we let  $\tilde{\xi}$  be the unique distribution on  $[0, L]$  such that

$$\langle \tilde{\xi}, f \rangle = \langle \xi, f \circ \sigma_L \rangle, \quad \forall f \in C^\infty([0, L]).$$

Since  $\sigma$  preserves the Lebesgue measure on  $[0, 1]$ , we deduce that  $\tilde{\xi}$  has the same law as  $\xi$ . Therefore, we can define the corresponding operator  $\tilde{\mathcal{H}} = -\partial_x^2 + \tilde{\xi}$  as well as the diffusions  $\tilde{X}_a$  subject to the noise  $\tilde{B}$ . Similarly, we can introduce the r.v.  $\tilde{a}_i < \tilde{a}'_i$  which approximate the  $k$  first eigenvalues of  $\tilde{\mathcal{H}}_L$ .

Take  $n \geq n_0$ . On the event  $\mathcal{E}(n, \varepsilon)$ , the following holds. For all  $a \geq a_0$  such that  $a \in M_L$ , the number of explosions of  $X_a$  (resp.  $\tilde{X}_a$ ) coincides with the sum of the number of explosions of  $X_a^j$  (resp.  $\tilde{X}_a^j$ ),  $j = 0, \dots, 2^n - 1$ . Furthermore, we have the identity

$$\tilde{X}_a^j(s_j^n L + t) = X_a^{2^{n-\sigma^{-1}(s_j^n)}}(\sigma_L^{-1}(s_j^n L) + t), \quad t \in [0, 2^{-n}L].$$

As a consequence, the number of explosions of  $X_a$  and  $\tilde{X}_a$  coincide and we deduce that  $a_i = \tilde{a}_i$  and  $a'_i = \tilde{a}'_i$ . This already ensures that the  $k$  first eigenvalues of  $\mathcal{H}_L$  and  $\tilde{\mathcal{H}}_L$  lie at a distance at most  $\varepsilon/\sqrt{a_L}$  from each other on the event  $\mathcal{E}(n, \varepsilon)$ . Additionally, the r.v.  $\tilde{U}_i/L$  falls into the same subinterval of length  $2^{-n}$  as  $\sigma(U_i/L)$ . Since  $\mathcal{E}(n, \varepsilon)$  has a probability of order  $1 - O(\varepsilon)$  for  $L$  and  $n$  large enough, we find:

$$\begin{aligned} & \mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{\tilde{U}_i/L \in [s_{j_i}^{n_0}, s_{j_i+1}^{n_0})} f(\tilde{\lambda}_1, \dots, \tilde{\lambda}_k) \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{\sigma(U_i/L) \in [s_{j_i}^{n_0}, s_{j_i+1}^{n_0})} f(\lambda_1, \dots, \lambda_k) \right] + O(\varepsilon) + O\left(\frac{\varepsilon}{\sqrt{a_L}}\right). \end{aligned}$$

On the other hand, since  $\mathcal{H}_L$  and  $\tilde{\mathcal{H}}_L$  have the same statistics, we immediately have

$$\begin{aligned} & \mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{\tilde{U}_i/L \in [s_{j_i}^{n_0}, s_{j_{i+1}}^{n_0}]} f(\tilde{\lambda}_1, \dots, \tilde{\lambda}_k) \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^k \mathbf{1}_{U_i/L \in [s_{j_i}^{n_0}, s_{j_{i+1}}^{n_0}]} f(\lambda_1, \dots, \lambda_k) \right]. \end{aligned}$$

By passing to the limit on a converging subsequence, we deduce that the identity (27) holds up to an error of order  $\varepsilon$ , which can be taken as small as desired, thus concluding the proof. □

### 4 Proofs of some first estimates on the diffusion $X_a$

#### 4.1 Invariant measure

For every given  $a > 0$ , the process  $(X_a(t), t \geq 0)$  admits a unique invariant (but not reversible) probability measure  $\mu_a(dx) = f_a(x)dx$  with

$$f_a(x) = \frac{2}{m(a)} \exp(-2V_a(x)) \int_{-\infty}^x \exp(2V_a(y)) dy. \tag{28}$$

To check this fact, one simply has to show that  $\mathcal{G}^* f_a = 0$ , where  $\mathcal{G}^*$  is the forward generator of the diffusion  $X_a$

$$\mathcal{G}^* f(x) = \frac{1}{2} f''(x) + (V'_a(x) f(x))'.$$

We also introduce the scale function associated with our diffusion:

$$S(x) = \int_{-\infty}^x \exp(2V_a(y)) dy = \int_{-\infty}^x \exp\left(-2ay + \frac{2}{3}y^3\right) dy.$$

If we let  $\mathbb{P}_x$  be the law of our diffusion starting from  $x \in (y, z)$ , then we have:

$$\mathbb{P}_x[\tau_y < \tau_z] = \frac{S(z) - S(x)}{S(z) - S(y)}. \tag{29}$$

In the next lemma, we establish some useful estimates on the invariant measure

**Lemma 4.1** *For all  $y(a)$  that goes to  $\infty$  as  $a \rightarrow \infty$  but is negligible compared to  $a^{3/4}$ , we have the bound*

$$\mu_a \left( \left[ \sqrt{a} - \frac{y(a)}{a^{1/4}}, \sqrt{a} + \frac{y(a)}{a^{1/4}} \right] \right) \geq 1 - e^{-y(a)^2},$$



uniformly over all  $a$  large enough. Furthermore, for all  $c \in (0, 1)$ , there exists  $\rho(c) > 0$  such that

$$\mu_a([(1 - c)\sqrt{a}, (1 + c)\sqrt{a}]) \geq 1 - e^{-\rho(c)a^{3/2}},$$

uniformly over all  $a$  large enough.

**Proof** Note that the density of the invariant measure writes:

$$\forall x \in \mathbf{R}, \quad f_a(x) = \frac{2}{m(a)} \exp(-2V(x))S(x). \tag{30}$$

Some rough estimates ensure that there exists  $\rho' > 0$  such that for all  $a$  large enough

$$\mu_a(\mathbf{R} \setminus [-\sqrt{a}/2, 3\sqrt{a}/2]) \leq e^{-\rho'a^{3/2}}.$$

Consequently, we can focus on the interval  $[-\sqrt{a}/2, 3\sqrt{a}/2]$ . Therein, the scale function is almost constant. Indeed, there exists  $\rho'' > 0$  such that for all  $x \in [-\sqrt{a}/2, 3\sqrt{a}/2]$ , we have  $S(x) = S(\sqrt{a})(1 + O(e^{-\rho''a^{3/2}}))$ . Moreover,  $S(\sqrt{a}) = \sqrt{\pi/2} a^{-1/4} \exp(4a^{3/2}/3)(1 + o(1))$ . Thus, using (30) we find for all  $x \in [-\sqrt{a}/2, 3\sqrt{a}/2]$ ,

$$f_a(x) = \sqrt{\frac{2}{\pi}} a^{1/4} \exp\left(-2\sqrt{a}(x - \sqrt{a})^2 \left(1 + \frac{(x - \sqrt{a})}{3\sqrt{a}}\right)\right) (1 + o(1)), \tag{31}$$

Recall that  $f_a$  integrates to 1. Therefore, to get the first bound of the statement it suffices to control the integral of  $f_a$  over  $[-\sqrt{a}/2, 3\sqrt{a}/2] \setminus [\sqrt{a} - \frac{y(a)}{a^{1/4}}, \sqrt{a} + \frac{y(a)}{a^{1/4}}]$ : this follows from a simple computation based on (31). The second bound is obtained similarly.  $\square$

### 4.2 Entrance and exit

In this paragraph, we will evaluate how much time the diffusion takes to go from  $+\infty$  down to various levels around  $\sqrt{a}$  (resp. the time it takes to explode from various levels near  $-\sqrt{a}$ ). Recall the definition of  $t_a$  and  $h_a$  given in (13).

**Lemma 4.2** (Entrance) *Take  $x(a) = \sqrt{a} + h_a/4$  and  $M = 2\sqrt{t_a} \ln \ln a$ . On the event  $\{\sup_{t \in [0, t_a]} |B(t)| < M\}$ , whose probability is at least  $1 - \exp(-(\ln \ln a)^2)$ , we have*

$$\sup_{t \in (0, \tau_{x(a)}(X))} |X(t) - \sqrt{a} \coth(\sqrt{at})| \lesssim M. \tag{32}$$

As a consequence, we get the following asymptotics for all  $u > 1$  and all  $c \geq 1/4$ :

$$\tau_{\sqrt{a}+c h_a}(X) = \frac{3}{8} \frac{\ln a}{\sqrt{a}} + \frac{1}{2\sqrt{a}} \ln \frac{2}{c \ln a} + o\left(\frac{1}{\sqrt{a}}\right), \tag{33}$$

$$\tau_{u\sqrt{a}}(X) = \frac{C(u)}{\sqrt{a}} + o\left(\frac{1}{\sqrt{a}}\right), \tag{34}$$

where  $u \mapsto C(u)$  is the reciprocal of  $x \mapsto \coth x$ .

Furthermore, we have the additional bound for all  $a$  large enough

$$\sup_{t \in (\tau_{x(a)}(X), \frac{3}{8}t_a]} |X(t) - \sqrt{a} \coth(\sqrt{at})| \leq 1. \tag{35}$$

**Proof** We adapt the arguments of the proof of [12, Proposition 2]. Let  $Z(t) = X(t) - B(t)$ . Necessarily,  $Z$  solves the ODE

$$dZ(t) = \left(a - Z(t)^2 \left(1 + \frac{B(t)}{Z(t)}\right)^2\right) dt, \quad Z(0) = +\infty.$$

We are therefore led to considering the ODE

$$dF(t) = \left(a - CF(t)^2\right) dt, \quad F(0) = +\infty,$$

with  $C$  being either  $C_1$  or  $C_2$  where

$$C_1 = \left(1 - \frac{M}{\ell(a) - M}\right)^2, \quad C_2 = \left(1 + \frac{M}{\ell(a) - M}\right)^2,$$

and  $\ell(a)$  is a value in  $[x(a), \infty)$  that will be chosen later on. The generic solution is given by  $F(t) = \sqrt{a/C} \coth(\sqrt{aC}t)$  for all  $t > 0$ . On the event  $\{\sup_{t \in [0, t_a]} |B(t)| < M\}$  and for  $a$  large enough, we have the following bound

$$F_2(t) \leq Z(t) \leq F_1(t), \quad t \in [0, \tau_{\ell(a)}(X) \wedge t_a].$$

Consequently the hitting time of  $\ell(a)$  by the diffusion  $X$  satisfies

$$\tau_{\ell(a)+M}(F_2) \wedge t_a \leq \tau_{\ell(a)}(X) \wedge t_a \leq \tau_{\ell(a)-M}(F_1) \wedge t_a.$$

We now derive the asymptotics of these upper and lower bounds. Let  $\ell_{\pm}(a)$  be  $\ell(a) \pm M$ .

When  $\ell(a) = u\sqrt{a}$ , we have:

$$\tau_{\ell_{\pm}(a)}(F) = \frac{C(u)}{\sqrt{a}} \left(1 + \mathcal{O}\left(\frac{M}{\ell_{\pm}(a)}\right)\right).$$

On the other hand, when  $\ell(a) = \sqrt{a} + c\frac{\ln a}{a^{1/4}}$ , the asymptotic expansion of  $\coth$  at infinity readily yields

$$\tau_{\ell_{\pm}(a)}(F) = -\frac{1}{2\sqrt{aC}} \ln \frac{1}{2} \left( \frac{\ell_{\pm}(a)}{\sqrt{a}} \sqrt{C} - 1 + o(e^{-2\sqrt{aC}\tau_{\ell_{\pm}(a)}(F)}) \right),$$

uniformly over all  $C$  in a neighborhood of 1. Since  $M \ll (\ell(a) - \sqrt{a})$ , a simple calculation shows that both  $\tau_{\ell(a)+M}(F_2)$  and  $\tau_{\ell(a)-M}(F_1)$  admit the following expansion

$$\frac{3}{8} \frac{\ln a}{\sqrt{a}} + \frac{1}{2\sqrt{a}} \ln \frac{2}{c \ln a} + o\left(\frac{1}{\sqrt{a}}\right),$$

as  $a \rightarrow \infty$ .

We have proven the asserted asymptotics on the hitting times: they ensure that  $\tau_{\ell(a)}(X) < t_a$  for  $a$  large enough, so that

$$F_2(t) - M \leq X(t) \leq F_1(t) + M, \quad t \in [0, \tau_{\ell(a)}(X)].$$

for all  $\ell(a) \in [x(a), \infty)$ . It is simple to check that  $F_2(t) - M \leq \sqrt{a} \coth(\sqrt{at}) \leq F_1(t) + M$  for all  $t \geq 0$ . Hence, for all  $\ell(a) \in [x(a), \infty)$

$$\begin{aligned} \sup_{t \in [\tau_{2\ell(a)}, \tau_{\ell(a)}]} |X(t) - \sqrt{a} \coth(\sqrt{at})| &\leq \sup_{t \in [\tau_{2\ell(a)+M}(F_2), \tau_{\ell(a)-M}(F_1)]} |(F_1(t)+M) - (F_2(t)-M)| \\ &\leq \sup_{t \in [\tau_{2\ell(a)+M}(F_2), \tau_{\ell(a)-M}(F_1)]} \left| \frac{F_1(t)}{F_2(t)} - 1 \right| F_2(t) + 2M. \end{aligned}$$

From the explicit expressions of  $F_1$  and  $F_2$ , we obtain:

$$\sup_{t > 0} \left| \frac{F_1(t)}{F_2(t)} - 1 \right| \lesssim \frac{M}{\ell(a)},$$

uniformly over all choices of  $\ell(a) \in [x(a), \infty)$ . Therefore,

$$\sup_{t \in [\tau_{2\ell(a)+M}(F_2), \tau_{\ell(a)-M}(F_1)]} \left| \frac{F_1(t)}{F_2(t)} - 1 \right| F_2(t) \lesssim M,$$

so that

$$\sup_{t \in [\tau_{2\ell(a)}, \tau_{\ell(a)}]} |X(t) - \sqrt{a} \coth(\sqrt{at})| \lesssim M,$$

uniformly over all choices of  $\ell(a) \in [x(a), \infty)$ . Patching together these estimates, we get (32).

To complete the proof of the lemma, it remains to control the diffusion on the interval  $[\tau_{x(a)}(X), (3/8)t_a]$  and on the event  $\{\sup_{t \in [0, t_a]} |B(t)| < M\}$ . To that end, we take  $\ell(a) = \sqrt{a}/2$  and, by the arguments at the beginning of the proof, we find

$$F_2(t) - M \leq X(t) \leq F_1(t) + M, \quad t \in [0, \tau_{\ell(a)}(X) \wedge t_a].$$

Since  $F_2((3/8)t_a) - M = \sqrt{a} - O(M)$  and since  $F_2$  is decreasing, we easily deduce that  $\tau_{\ell(a)+M}(F_2) > (3/8)t_a$  and therefore  $\tau_{\ell(a)}(X) > t_a$ . This yields (35).  $\square$

We have an analogous result right before the explosion time, we keep the notations from the previous lemma.

**Lemma 4.3** (Explosion) *On the event  $\{\sup_{t \in [0, t_a]} |B(\tau_{-x(a)}(X) + t) - B(\tau_{-x(a)}(X))| < M\}$ , whose probability is at least  $1 - \exp(-(\ln \ln a)^2)$ , we have:*

$$\sup_{t \in [\tau_{-x(a)}(X), \tau_{-\infty}(X)]} \left| X(t) + \sqrt{a} \coth(\sqrt{a}(\tau_{-\infty}(X) - t)) \right| \lesssim M.$$

Therefore, for  $u > 1$  and  $c \geq 1/4$ , as  $a \rightarrow \infty$

$$\begin{aligned} \tau_{-\infty}(X) - \tau_{-\sqrt{a}-c h_a}(X) &= \frac{3}{8} \frac{\ln a}{\sqrt{a}} + \frac{1}{2\sqrt{a}} \ln \frac{2}{c \ln a} + o\left(\frac{1}{\sqrt{a}}\right), \\ \tau_{-\infty}(X) - \tau_{-u\sqrt{a}}(X) &= \frac{C(u)}{\sqrt{a}} + o\left(\frac{1}{\sqrt{a}}\right). \end{aligned}$$

**Proof** The proof is essentially the same as that of Lemma 4.2 so we do not provide the details. □

**Proof of Proposition 2.5** By Lemmas 4.2 and 4.3, the trajectory  $(X(t), t \in [0, \zeta(1)])$  satisfies the two requirements of the event  $\mathcal{D}_a^N$  with a probability at least  $1 - \exp(-(\ln \ln a)^2)$  uniformly over all  $a$  large enough. Since the random paths  $(X(t + \zeta(k)), t \in [0, \zeta(k + 1) - \zeta(k)])$ ,  $k \in \{0, \dots, N\}$  are i.i.d, we easily deduce the statement of Proposition 2.5. □

### 4.3 Oscillations

The goal of this subsection is to prove Proposition 2.6. From now on,  $Y$  is taken to be a solution of (7) starting from the stationary measure  $\mu_a$ . The key step consists in showing that  $Y$  spends most of its time near  $\sqrt{a}$ . We introduce the notations:

$$\begin{aligned} \tau_{\leq \ell}(f) &:= \inf\{t \geq 0 : f(t) \leq \ell\}, & \tau_{\geq \ell}(f) &:= \inf\{t \geq 0 : f(t) \geq \ell\}, \\ \tau_{-\infty}^k(f) &:= \inf\{t \geq \tau_{-\infty}^{k-1}(f) : f(t) = -\infty\}, & \tau_{-\infty}^0(f) &:= 0. \end{aligned}$$

For  $r > 0$ , we define the following interval around the bottom of the well  $\sqrt{a}$ :

$$I_a(r) := [\sqrt{a} - r h_a, \sqrt{a} + r h_a].$$

We start with a simple estimate on the exit times of  $X$ .

**Lemma 4.4** *For any  $T > 0$  and any  $0 < d < D < 2\sqrt{a}$  we have*

$$\mathbb{P}(\exists t \in [0, T], X(t) \notin [\sqrt{a} - D, \sqrt{a} + D] \mid X(0) \in [\sqrt{a} - d, \sqrt{a} + d]) \leq \frac{8\sqrt{T}}{D-d} e^{-\frac{(D-d)^2}{2T}}.$$

**Proof** We let  $A$  be the reflected Brownian motion starting from  $d$ , obtained as the solution of the following Skorohod problem

$$dA(t) = dB(t) + d\ell(t), \quad \int_{t \geq 0} A(t)d\ell(t) = 0, \quad A(0) = d.$$

For  $x \in [\sqrt{a} - d, \sqrt{a} + d]$ , consider the diffusion  $X$  starting from  $x$ . We claim that  $A(t) - (X(t) - \sqrt{a}) \geq 0$  for all  $t \geq 0$ . Indeed, for any time  $t \geq 0$  at which  $A(t) = X(t) - \sqrt{a}$ , we have  $d(A(t) - (X(t) - \sqrt{a})) > 0$  since either  $A(t) = X(t) - \sqrt{a} = 0$  and then  $d\ell_t > 0$  and  $-V'(X(t)) = 0$ , or  $A(t) = X(t) - \sqrt{a} > 0$  and then  $d\ell(t) = 0$  and  $-V'(X(t)) < 0$ . Since  $A(t)$  has the same law as  $|B(t) + d|$  and since the supremum (resp. the infimum) of  $B$  on  $[0, T]$  has the same law as  $|B(T)|$  (resp.  $-|B(T)|$ ) we obtain

$$\begin{aligned} \mathbb{P}(\sup_{t \in [0, T]} X(t) > \sqrt{a} + D | X(0) = x) &\leq \mathbb{P}(\sup_{t \in [0, T]} A(t) > D) \leq 2\mathbb{P}(|B(T)| > D - d) \\ &\leq \frac{4\sqrt{T}}{D - d} e^{-\frac{(D-d)^2}{2T}}. \end{aligned}$$

The same argument applies to bound  $X$  from below, thus concluding the proof. □

**Lemma 4.5** (Oscillations of  $Y$ ) *There exists  $c > 0$  such that the probability that*

$$\int_0^t Y(s)ds \in I_a(1/2), \quad \forall t \in [0, \tau_{\leq -3\sqrt{a}}],$$

*is larger than  $1 - \exp(-c(\ln a)^2)$  uniformly over all  $a$  large enough.*

**Remark 4.6** We could improve the bounds by choosing an interval of size  $C\sqrt{\ln a}/a^{1/4}$  around  $\sqrt{a}$  at the cost of decreasing the lower bound on the probability.

**Proof** Let  $L_a = \exp(b a^{3/2})$  for some large enough  $b > 0$ . We first control the integral of the statement for all  $t \in [1/m(a), \tau_{\leq -3\sqrt{a}}(Y) \wedge \tau_{\geq 4\sqrt{a}}(Y) \wedge L_a]$ . For any such  $t$ , we have:

$$\int_0^t Y(s)ds = \int_0^t Y(s)\mathbf{1}_{\{Y(s) \in I_a(1/4)\}}ds + \int_0^t Y(s)\mathbf{1}_{\{Y(s) \notin I_a(1/4)\}}ds.$$

For  $k \in \mathbb{N}$ , we set

$$E_k := \left\{ \int_0^{2^{-k} L_a} \mathbf{1}_{\{Y(s) \notin I_a(1/4)\}}ds \geq \frac{1}{a} \right\}.$$

Using the Markov inequality, we find  $\mathbb{P}(E_k) \leq a \mu_a(I_a(1/4)^c)$ . Let  $k^*$  be the smallest integer such that  $2^{-k^*} L_a < 1/m(a)$ . Observe that  $k^*$  is of order  $a^{3/2}$  thanks to our

hypothesis on the growth of  $L_a$  so that thanks to Lemma 4.1,

$$\mathbb{P}\left[\bigcup_{k=0}^{k^*} E_k\right] \leq \sum_{k \leq k^*} \mathbb{P}[E_k] \lesssim a^{5/2} e^{-(\ln a)^2/16}.$$

On the event  $E_k^c$ , we get for all  $t \in [2^{-k-1}L_a, 2^{-k}L_a]$  such that  $t \leq \tau_{\leq -3\sqrt{a}}(Y) \wedge \tau_{\geq 4\sqrt{a}}(Y)$ , and all  $a$  large enough, the following upper bound

$$\begin{aligned} \int_0^t Y(s)ds &\leq (\sqrt{a} + h_a/4)t + 4\sqrt{a} \int_0^t 1_{\{Y(s) \notin I_a(1/4)\}} ds \\ &\leq (\sqrt{a} + h_a/2)t, \end{aligned}$$

and the following lower bound

$$\int_0^t Y(s)ds \geq (\sqrt{a} - h_a/2)t.$$

To conclude the proof, we show that with large probability  $Y$  does not hit  $[4\sqrt{a}, \infty)$  before  $(-\infty, -3\sqrt{a}]$ , nor exit  $I_a(1/4)$  within a time  $1/m(a)$ , and that  $Y$  hits  $(-\infty, -3\sqrt{a}]$  before time  $L_a$ . Regarding the first claim, we have for all  $a > 0$  large enough

$$\begin{aligned} \mathbb{P}[\tau_{\geq 4\sqrt{a}}(Y) < \tau_{\leq -3\sqrt{a}}(Y)] &= \int_{x \in \mathbf{R}} \mathbb{P}_x[\tau_{4\sqrt{a}} < \tau_{-3\sqrt{a}}] \mu_a(dx) \\ &\leq \mu_a([-2\sqrt{a}, 2\sqrt{a}]) \mathbb{P}_{2\sqrt{a}}[\tau_{4\sqrt{a}} < \tau_{-3\sqrt{a}}] + O(e^{-\rho'a^{3/2}}) \\ &\leq e^{-\rho'a^{3/2}}, \end{aligned}$$

for some positive  $\rho$ , using (29) and some simple estimates on the scale function. The second claim follows from Lemma 4.4 combined with the bounds of Lemma 4.1 on the invariant measure. The last claim follows from  $\mathbb{P}[\tau_{\leq -3\sqrt{a}}(Y) > L_a] \leq \mathbb{P}_{+\infty}[\tau_{-\infty}(X)/m(a) > L_a/m(a)]$  and Markov's inequality together with (10).  $\square$

The following lemma shows that, after  $X$  has come down from infinity,  $X$  and  $Y$  stay very close to each other until  $Y$  starts following its deterministic path to  $-\infty$ .

**Lemma 4.7** *There exist  $b, C > 0$  such that with probability at least  $1 - e^{-b(\ln \ln a)^2}$ , we have:*

- (i)  $\tau_{\leq -3\sqrt{a}}(Y) > (3/8)t_a$  and  $\tau_{\leq -2\sqrt{a}}(X) > (3/8)t_a$
- (ii)  $|X(t) - Y(t)| \leq h_a/2$  for all  $t \in [(3/8)t_a, \tau_{\leq -3\sqrt{a}}(Y)]$ ,
- (iii)  $|\tau_{-\infty}(X) - \tau_{-\infty}(Y)| < \frac{C}{\sqrt{a}}$
- (iv)  $\int_{(3/8)t_a}^t X(s)ds \in I_a(1)$  for all  $t \in [(3/8)t_a, \tau_{-2\sqrt{a}}(X)]$ .

**Proof** By Lemmas 4.4 and 4.1, it is simple to establish the first bound of (i) with  $-3\sqrt{a}$  replaced by  $-2\sqrt{a}$ . Monotonicity then easily yields the second bound of (i). Recall that  $X(t) \geq Y(t)$  for all  $t \in [0, \tau_{-\infty}(Y)) \supset [0, \tau_{\leq -3\sqrt{a}}(Y))$ . The difference  $Z = X - Y$  satisfies the ODE:

$$dZ(t) = -Z(t)(X(t) + Y(t)), \quad t \geq 0.$$

By Lemma 4.2, there exists  $z > 0$  such that if we set  $t_1 := (3/8)t_a + \ln(z/\ln a)/(2\sqrt{a})$  then with probability greater than  $1 - \exp(-(\ln \ln a)^2)$ , we have  $X(t_1) \in I_a(1/3)$ . By the first estimate on the invariant measure in Lemma 4.1, we deduce that there exists  $c' > 0$  such that for all  $a$  large enough

$$\mathbb{P}[Z(t_1) \leq h_a/2] \geq \mathbb{P}[X(t_1) \in I_a(1/3), Y(t_1) \in I_a(1/6)] \geq 1 - e^{-c'(\ln \ln a)^2}.$$

By Lemma 4.5 applied from time  $t_1$  (recall that  $Y$  is stationary) we have with probability at least  $1 - e^{-c(\ln a)^2}$

$$\int_{t_1}^t Y(s)ds \in I_a(1/2), \quad \forall t \in [t_1, \tau'_{\leq -3\sqrt{a}}(Y)],$$

where  $\tau'_{\leq -3\sqrt{a}}(Y) := \inf\{t \geq t_1 : Y(t) \leq -3\sqrt{a}\}$ . By (i), we can assume that  $\tau'_{\leq -3\sqrt{a}}(Y) = \tau_{\leq -3\sqrt{a}}(Y)$ . Henceforth, there exists  $b' > 0$  such that the probability that for all  $t \in [t_1, \tau_{\leq -3\sqrt{a}}(Y)]$

$$Z(t) = Z(t_1) \exp\left(-\int_{t_1}^t (X(s) + Y(s))ds\right) \leq Z(t_1) \exp\left(-2 \int_{t_1}^t Y(s)ds\right) \leq \frac{h_a}{2} e^{-2(\sqrt{a} - h_a/2)(t - t_1)},$$

is larger than  $1 - \exp(-b'(\ln \ln a)^2)$  for all  $a$  large enough. This proves (ii). Applying Lemma 4.3, we get (iii). Applying Lemma 4.5 from time  $(3/8)t_a$  and using (ii), we easily get (iv). □

More can be said about  $X - Y$ . Since  $Y(\tau_{-\infty}(Y) + t), t \geq 0$  has the same distribution as  $X$ , we deduce that it satisfies the estimates (33) and (iv) of Lemma 4.7. Since with large probability  $\tau_{-\infty}^1(X) - \tau_{-\infty}^1(Y)$  is at most  $C/\sqrt{a}$ , a simple iteration of the proof of Lemma 4.7 allows to prove that with probability at least  $1 - 2e^{-b(\ln \ln a)^2}$ , we have  $\tau_{-3\sqrt{a}}^2(Y) > \tau_{-\infty}(X) + (3/8)t_a, \tau_{-2\sqrt{a}}^2(X) > \tau_{-\infty}(X) + (3/8)t_a$  as well as

$$|X(t) - Y(t)| \leq h_a/2, \quad \forall t \in [\tau_{-\infty}(X) + (3/8)t_a, \tau_{-3\sqrt{a}}^2(Y)],$$

and  $\tau_{-\infty}^2(X) - \tau_{-\infty}^2(Y) \leq C/\sqrt{a}$ . Iterating this, we get the following result.

**Corollary 4.8** *There exist  $b, C > 0$  such that for any  $N_a \geq 1$  with probability at least  $1 - N_a e^{-b(\ln \ln a)^2}$  the following holds for all  $a$  large enough. For all  $k \in \{1, \dots, N_a\}$*

we have

$$\begin{aligned}\tau_{-3\sqrt{a}}^k(Y) &> \tau_{-\infty}^{k-1}(X) + \frac{3}{8}t_a, & \tau_{-2\sqrt{a}}^k(X) &> \tau_{-\infty}^{k-1}(X) + (3/8)t_a, \\ |X(t) - Y(t)| &\leq h_a/2, & \text{for all } t &\in [\tau_{-\infty}^{k-1}(X) + \frac{3}{8}t_a, \tau_{-3\sqrt{a}}^k(Y)], \\ |\tau_{-\infty}^k(X) - \tau_{-\infty}^k(Y)| &< \frac{C}{\sqrt{a}}.\end{aligned}$$

where  $\tau_{-3\sqrt{a}}^k(Y) := \inf\{t \geq \tau_{-\infty}^{k-1}(Y) : Y(t) = -3\sqrt{a}\}$ .

We conclude this section with the proof of Proposition 2.6.

**Proof of Proposition 2.6** Using Corollary 4.8, with probability greater than  $1 - e^{-b(\ln \ln a)^2}$ , we have:

$$\begin{aligned}\tau_{-3\sqrt{a}}^k(Y) &> \tau_{-\infty}^{k-1}(X) + \frac{3}{8}t_a, & \tau_{-2\sqrt{a}}^k(X) &> \tau_{-\infty}^{k-1}(X) + (3/8)t_a, \\ |X(t) - Y(t)| &\leq h_a/2, & \forall t &\in \left[ \zeta_a(k-1) + \frac{3}{8}t_a, \tau_{-3\sqrt{a}}^k(Y) \right], & (36) \\ |\zeta_a(k) - \tau_{-\infty}^k(Y)| &< \frac{C}{\sqrt{a}}.\end{aligned}$$

Let  $Z(t) = X^{t_0}(t) - Y(t)$ ,  $\tau := \inf\{t \geq t_0 : Y(t) \leq -3\sqrt{a}\}$  and  $\tau' := \inf\{t \geq t_0 : Y(t) = -\infty\}$  then the same arguments as in the proof of Lemma 4.7 ensure that with probability at least  $1 - e^{-b(\ln \ln a)^2}$  we have  $\tau > t_0 + (3/8)t_a$  as well as

$$\begin{aligned}Z(t) &\leq h_a/2, & \forall t &\in [t_0 + (3/8)t_a, \tau], \\ |\tau_{-\infty}(X^{t_0}) - \tau'| &< \frac{C}{\sqrt{a}}.\end{aligned}$$

By Lemma 4.4,  $Y(t) > 0$  for all  $t \in [(t_0 - (3/8)t_a)_+, t_0 + (3/8)t_a]$  with large probability: we can therefore work on this event from now on. If  $t_0 \in [\zeta_a(k-1), \zeta_a(k))$ , then by (36) and the positivity of  $Y$  on  $[(t_0 - (3/8)t_a)_+, t_0 + (3/8)t_a]$  we have  $t_0 + (3/8)t_a < \tau_{-\infty}^k(Y)$ . Therefore  $\tau' = \tau_{-\infty}^k(Y)$  and  $\tau = \tau_{-3\sqrt{a}}^k(Y)$ . Putting everything together, we get the statement of the proposition.  $\square$

## 5 Crossing of the well

This section is devoted to the proof of Propositions 3.3, 3.5, 3.6 and of part (b) of Proposition 3.9. This is achieved through a series of technical lemmas. The proof of Proposition 3.3 (i) and (ii) can be found at the end of Sect. 5.1. The proof of Proposition 3.5 is in Sect. 5.2. The proofs of Proposition 3.3 (iii) and Proposition 3.6 can be found in Sect. 5.3. Finally, the proof of part (b) of Proposition 3.9 is presented in Sects. 5.4 and 5.5.



### 5.1 Bounds up to time $\theta_a$

In the sequel, we take the following parameters:

$$T = \frac{3}{4}t_a, \quad \delta = \frac{(\ln a)^2}{a^{1/4}},$$

and we study the diffusion  $X$  when it crosses the region  $[-\sqrt{a}, \sqrt{a}]$ .

Let us introduce the diffusion:

$$dH(t) = (-a + H^2(t))dt + dB(t),$$

whose drift is the opposite of the drift of  $X$ . This diffusion turns out to be a good approximation of the diffusion  $X$  starting below  $\sqrt{a}$  and conditioned to hit  $-\sqrt{a}$  before  $\sqrt{a}$  (see also [12] where similar techniques were used). This can be stated precisely thanks to Girsanov’s theorem.

We will denote by  $\mathbf{P}_x$  the law of this diffusion starting from  $x$ , while  $\mathbb{P}_x$  will denote the law of  $X_a$  starting from  $x$ .

**Lemma 5.1** (Diffusion from  $\sqrt{a} - \delta$  to  $-\sqrt{a} + \delta$ ) *Recall that  $\tau_0$  denotes the first hitting time of 0 by  $X$ . For any  $c > 0$ , there exists  $C > 0$  such that for all  $a$  large enough, we have:*

$$\mathbb{P}_{\sqrt{a}-\delta}[E(C) \mid \tau_{-\sqrt{a}+\delta} < \tau_{\sqrt{a}-\delta/2} \wedge T] \leq a^{-c},$$

where  $E(C)$  is the following event

$$E(C) := \left\{ \sup_{t \in [0, \tau_{-\sqrt{a}+\delta}]} |X(t) - \sqrt{a} \tanh(-\sqrt{a}(t - \tau_0))| \geq C \frac{\sqrt{a}}{\ln a} \right\} \\ \cup \left\{ \left| \tau_0 - \frac{3}{8}t_a \right| \geq C \frac{\ln \ln a}{\sqrt{a}} \right\} \cup \left\{ \left| \tau_{-\sqrt{a}+\delta} - \tau_0 - \frac{3}{8}t_a \right| \geq C \frac{\ln \ln a}{\sqrt{a}} \right\}.$$

Notice that, eventually, the stopping time  $\tau_0$  will correspond (up to a negligible error) to the r.v.  $\nu_a$  in the context of Proposition 3.3.

**Proof** To simplify the notations, we set

$$\tau_+ := \tau_{\sqrt{a}-\delta/2}, \quad \tau_- := \tau_{-\sqrt{a}+\delta}.$$

If we stop the diffusions at time  $T \wedge \tau_- \wedge \tau_+$ , then Girsanov’s Theorem [35, Th.VIII.1.7] ensures that the Radon-Nikodym derivative of the law of  $X_a$  w.r.t. the law of  $H$  up to time  $t$  is given by the exponential of  $G_t(H) := 2 \int_0^t (a - H(s)^2)dH(s)$ . Applying Itô’s formula to  $V(H)$ , a simple calculation yields:

$$G_t(H) = 2(V_a(H(0)) - V_a(H(t))) + 2 \int_0^t H(s)ds.$$

We thus get:

$$\begin{aligned} \frac{\mathbb{P}_{\sqrt{a}-\delta}[E ; \tau_- < \tau_+ \wedge T]}{\mathbb{P}_{\sqrt{a}-\delta}[\tau_- < \tau_+ \wedge T]} &= \frac{\mathbf{E}_{\sqrt{a}-\delta}[E ; \tau_- < \tau_+ \wedge T ; \exp(G_{\tau_-}(H))]}{\mathbf{E}_{\sqrt{a}-\delta}[\tau_- < \tau_+ \wedge T ; \exp(G_{\tau_-}(H))]} \\ &= \frac{\mathbf{E}_{\sqrt{a}-\delta}[E ; \tau_- < \tau_+ \wedge T ; \exp(2 \int_0^{\tau_-} H(t)dt)]}{\mathbf{E}_{\sqrt{a}-\delta}[\tau_- < \tau_+ \wedge T ; \exp(2 \int_0^{\tau_-} H(t)dt)]} \\ &\leq \exp(4\sqrt{a} T) \frac{\mathbf{P}_{\sqrt{a}-\delta}[E ; \tau_- < \tau_+ \wedge T]}{\mathbf{P}_{\sqrt{a}-\delta}[\tau_- < \tau_+ \wedge T]}. \end{aligned} \tag{37}$$

We now study  $H$  in the region  $[-\sqrt{a} + \delta, \sqrt{a} - \delta/2]$ . Notice that  $Z := H - B$  satisfies

$$dZ(t) = (-a + (Z(t) + B(t))^2)dt.$$

Set  $M := c_0 \ln a/a^{1/4}$  and assume that  $\sup_{t \in [0, T]} |B(t)| \leq M$  holds true. Then, if  $|H(t)| \in [\sqrt{a}/2, \sqrt{a}]$  we have

$$\begin{aligned} -a + (Z(t) + B(t))^2 &\leq -a + Z(t)^2 \left(1 + \frac{M}{\sqrt{a}/2 - M}\right)^2 \\ &\leq -a + Z(t)^2 \left(1 + \frac{4M}{\sqrt{a}}\right)^2, \end{aligned} \tag{38}$$

for  $a$  large enough. On the other hand, if  $|H(t)| \in [0, \sqrt{a}/2]$ , we have

$$-a + (Z(t) + B(t))^2 \leq (-a + Z(t)^2) \left(1 - \frac{4M}{\sqrt{a}}\right). \tag{39}$$

Therefore, whenever  $\sup_{t \in [0, T]} |B(t)| \leq M$ , and  $|H(t)| \leq \sqrt{a}$ , we get thanks to (38) and (39):

$$-a + (Z(t) + B(t))^2 \leq -a \left(1 - \frac{4M}{\sqrt{a}}\right) + Z(t)^2 \left(1 + \frac{4M}{\sqrt{a}}\right)^2.$$

Similarly, we have the lower bound:

$$-a + (Z(t) + B(t))^2 \geq -a \left(1 + \frac{4M}{\sqrt{a}}\right) + Z(t)^2 \left(1 - \frac{4M}{\sqrt{a}}\right)^2.$$

We deduce that when  $\sup_{t \in [0, T]} |B(t)| \leq M$  and as long as  $|H(t)| \leq \sqrt{a}$  we have  $Z_-(t) \leq Z(t) \leq Z_+(t)$  with

$$dZ_+(t) = \left(-a \left(1 - \frac{4M}{\sqrt{a}}\right) + Z_+(t)^2 \left(1 + \frac{4M}{\sqrt{a}}\right)^2\right) dt, \quad Z_+(0) = \sqrt{a} - \delta,$$

$$dZ_{-}(t) = \left( -a \left( 1 + \frac{4M}{\sqrt{a}} \right) + Z_{-}(t)^2 \left( 1 - \frac{4M}{\sqrt{a}} \right)^2 \right) dt, \quad Z_{-}(0) = \sqrt{a} - \delta.$$

Denote by  $\kappa = 4M/\sqrt{a}$ . Those last equations have explicit solutions given by

$$\begin{aligned} Z_{+}(t) &= -\frac{\sqrt{a}(1-\kappa)^{1/2}}{1+\kappa} \tanh(\sqrt{a}(1-\kappa)^{1/2}(1+\kappa)(t-A_{+})), \\ Z_{-}(t) &= -\frac{\sqrt{a}(1+\kappa)^{1/2}}{1-\kappa} \tanh(\sqrt{a}(1+\kappa)^{1/2}(1-\kappa)(t-A_{-})), \end{aligned}$$

where

$$A_{+} = \frac{1}{2\sqrt{a}} \left( \ln \frac{2\sqrt{a}}{\delta} + \frac{3\sqrt{a}}{2\delta} \kappa + o\left(\frac{1}{\ln a}\right) \right), \quad A_{-} = \frac{1}{2\sqrt{a}} \left( \ln \frac{2\sqrt{a}}{\delta} - \frac{3\sqrt{a}}{2\delta} \kappa + o\left(\frac{1}{\ln a}\right) \right).$$

Note that

$$A_{\pm} = \frac{3 \ln a}{8 \sqrt{a}} - \frac{\ln \ln a}{\sqrt{a}} + \frac{\ln(2)}{2\sqrt{a}} \pm \frac{3c_0}{\sqrt{a} \ln a} (1 + o(1)).$$

From there, simple calculations show that  $Z_{+}(\cdot) + M$  stays below  $\sqrt{a} - \delta/2$  and passes below  $-\sqrt{a} + \delta$  before time  $T$ , while  $Z_{-}(\cdot) - M$  is above  $-\sqrt{a} + \delta$  at time  $T - 3 \ln \ln a / \sqrt{a}$ . Furthermore,  $Z_{+}(\cdot) + M$  and  $Z_{-}(\cdot) - M$  always stay at a distance of order  $\sqrt{a} / \ln a$  from one another. Their respective crossing times of 0 lie at  $(3/8)t_a + O(\ln \ln a / \sqrt{a})$  and are at a distance of order  $1/(\ln a \sqrt{a})$  from one another. Recall that  $Z_{-}(t) - M \leq H(t) \leq Z_{+}(t) + M$  for  $t \in [0, T]$  when  $\sup_{t \in [0, T]} |B(t)| \leq M$  and  $\sup_{t \in [0, T]} |H(t)| \leq \sqrt{a}$  so that all the hitting times of 0 by  $H$  are located near the crossing times of 0 of the latter two curves. Hence, there exists  $C > 0$  such that on the event  $\sup_{t \in [0, T]} |B(t)| \leq M$  we have

$$\begin{aligned} \sup_{t \in [0, \tau_{-\sqrt{a}+\delta}]} |H(t) - \sqrt{a} \tanh(-\sqrt{a}(t - \tau_0))| &< C \frac{\sqrt{a}}{\ln a}, \\ |\tau_0 - \frac{3}{8}t_a| &< C \frac{\ln \ln a}{\sqrt{a}}, \quad |\tau_{-\sqrt{a}+\delta} - \tau_0 - \frac{3}{8}t_a| < C \frac{\ln \ln a}{\sqrt{a}}. \end{aligned}$$

Since for  $a$  large enough we have  $\mathbf{P}(\sup_{t \in [0, T]} |B(t)| \geq M) \lesssim \exp(-M^2/(2T)) = a^{-\frac{2}{3}c_0^2}$ , we get

$$(37) \leq \exp(4\sqrt{a} T) \frac{\mathbf{P}[\sup_{t \in [0, T]} |B(t)| \geq M]}{\mathbf{P}[\sup_{t \in [0, T]} |B(t)| \leq M]} \lesssim a^{3-\frac{2}{3}c_0^2},$$

for all  $a$  large enough. This concludes the proof. □

**Lemma 5.2** (Time needed to go from  $\sqrt{a} - \delta$  to  $-\sqrt{a} + \delta$ ) *For all  $r > 0$  uniformly over all large  $a$  we have:*

$$\mathbb{P}_{\sqrt{a}-\delta} \left[ \tau_{-\sqrt{a}+\delta} > T \mid \tau_{-\sqrt{a}+\delta} < \tau_{\sqrt{a}-\delta/2} \right] \lesssim a^{-r}.$$

**Proof** Fix  $r > 0$ . We keep the notation of the preceding proof. Since

$$\mathbb{P}_{\sqrt{a}-\delta} \left( \tau_- \geq T \mid \tau_- < \tau_+ \right) = \lim_{T' \rightarrow \infty} \frac{\mathbb{P}_{\sqrt{a}-\delta} (T \leq \tau_- < T' \wedge \tau_+)}{\mathbb{P}_{\sqrt{a}-\delta} (\tau_- < T' \wedge \tau_+)},$$

and since  $\mathbb{P}_{\sqrt{a}-\delta} (\tau_- < T' \wedge \tau_+) \geq \mathbb{P}_{\sqrt{a}-\delta} (\tau_- < T \wedge \tau_+)$  for  $T' > T$ , the lemma will follow if we prove that for  $c$  large enough we have the following bound

$$\sup_{T' > T} \frac{\mathbb{P}_{\sqrt{a}-\delta} (T \leq \tau_- < T' \wedge \tau_+)}{\mathbb{P}_{\sqrt{a}-\delta} (\tau_- < T \wedge \tau_+)} \lesssim a^{-r}. \tag{40}$$

By Girsanov’s Theorem applied to the diffusions stopped at time  $T' \wedge \tau_- \wedge \tau_+$ , it suffices to show that

$$\sup_{T' > T} \frac{\mathbf{E}_{\sqrt{a}-\delta} \left[ T \leq \tau_- < T' \wedge \tau_+ ; e^{2 \int_0^{\tau_-} H(t) dt} \right]}{\mathbf{E}_{\sqrt{a}-\delta} \left[ \tau_- < T \wedge \tau_+ ; e^{2 \int_0^{\tau_-} H(t) dt} \right]} \lesssim a^{-r}.$$

We let  $n' := \lfloor T'/T \rfloor$  and we set for every  $n \geq 0$  and every  $x \in \mathbb{R}$

$$A_x(n) := \mathbf{E}_x \left[ \tau_- < \tau_+ ; \tau_- \in [nT, (n+1)T) ; e^{2 \int_0^{\tau_-} H(t) dt} \right].$$

We have

$$\mathbf{E}_{\sqrt{a}-\delta} \left[ T \leq \tau_- < T' \wedge \tau_+ ; e^{2 \int_0^{\tau_-} H(t) dt} \right] \leq \sum_{n=1}^{n'} A_{\sqrt{a}-\delta}(n).$$

Applying the Markov property at time  $T$ , and bounding  $H$  by  $\sqrt{a}$  on the interval  $[0, T)$ , we get for all  $n \geq 1$  and  $x \in (-\sqrt{a} + \delta, \sqrt{a} - \delta/2)$ ,

$$\begin{aligned} A_x(n) &= \mathbf{E}_x \left[ T \leq \tau_- < \tau_+ ; e^{2 \int_0^T H(s) ds} ; A_{H(T)}(n-1) \right] \\ &\leq e^{2\sqrt{a}T} \mathbf{P}_x \left[ T \leq \tau_- \right] \sup_{y \in (-\sqrt{a}+\delta, \sqrt{a}-\delta/2)} A_y(n-1). \end{aligned}$$

On the other hand, we obviously have the bound

$$A_y(0) \leq e^{2\sqrt{a}T} \mathbf{P}_y \left[ \tau_- < T \wedge \tau_+ \right] \leq e^{2\sqrt{a}T}.$$

Thus, a simple recursion for the first bound and the monotonicity of the process for the second bound, yields

$$\begin{aligned}
 A_{\sqrt{a}-\delta}(n) &\leq e^{(n+1)2\sqrt{a}T} \sup_{y \in (-\sqrt{a}+\delta, \sqrt{a}-\delta/2)} \mathbf{P}_y[T < \tau_-]^n \\
 &\leq e^{(n+1)2\sqrt{a}T} \mathbf{P}_{\sqrt{a}-\delta/2}[T < \tau_-]^n.
 \end{aligned}$$

By the computations made in the proof of Lemma 5.1, we have

$$\mathbf{P}_{\sqrt{a}-\delta/2}[T < \tau_-] \lesssim a^{-c'},$$

for any  $c' > 0$ , as well as

$$\mathbf{E}_{\sqrt{a}-\delta}[\tau_- < T \wedge \tau_+; e^{2 \int_0^{\tau_-} H(t)dt}] \geq (1 - O(a^{-c'}))e^{-2\sqrt{a}T} \geq \frac{1}{2}e^{-2\sqrt{a}T}.$$

Putting everything together, we get that for all  $T' > T$ ,

$$\mathbb{P}_{\sqrt{a}-\delta}[\tau_- < T' \mid \tau_- < \tau_+] \lesssim e^{4\sqrt{a}T} \sum_{n \geq 1} (e^{2\sqrt{a}T} a^{-c'})^n,$$

which is bounded by a term of order  $a^{-r}$ , for any given  $r > 0$ , provided  $c'$  is large enough. □

We now control the portion of trajectory from  $-\sqrt{a} + \delta$  to  $-\sqrt{a}$ .

**Lemma 5.3** *For all  $C > 1$  we have*

$$\mathbb{P}_{-\sqrt{a}+\delta} \left[ \tau_{-\sqrt{a}} > C \frac{\ln \ln a}{\sqrt{a}} \wedge \tau_{-\sqrt{a}+2\delta} \mid \tau_{-\sqrt{a}} < \tau_{\sqrt{a}-\delta/2} \right] \lesssim (\ln a)^{2-2C}.$$

**Proof** We set  $S = C \frac{\ln \ln a}{\sqrt{a}}$ ,  $\tau_- := \tau_{-\sqrt{a}}$ ,  $\tau_+ := \tau_{-\sqrt{a}+2\delta}$  and  $\tau_{++} := \tau_{\sqrt{a}-\delta/2}$ . For convenience, we also set  $I(a) := e^{2(V(-\sqrt{a}+\delta)-V(-\sqrt{a}))}$  as this term will pop up in many equations below. We are going to show that we have

$$\mathbb{P}_{-\sqrt{a}+\delta} [\tau_- < S \wedge \tau_+] \gtrsim I(a)e^{-2\sqrt{a}S},$$

as well as

$$\mathbb{P}_{-\sqrt{a}+\delta} [S \wedge \tau_+ \leq \tau_- < S' \wedge \tau_{++}] \lesssim I(a)e^{-2\sqrt{a}S} (\ln a)^{2-2C},$$

uniformly over all  $S' > 0$  and all  $a$  large enough. These two bounds yield the statement of the lemma.

To prove these bounds we apply Girsanov’s Theorem to the diffusions stopped at time  $S' \wedge \tau_- \wedge \tau_{++}$ . This will allow us to approximate our process by an Ornstein-Uhlenbeck process and use standard estimates about OU. Regarding the first term, we have:

$$\begin{aligned} \mathbb{P}_{-\sqrt{a}+\delta}[\tau_- < S \wedge \tau_+] &= I(a)\mathbf{E}_{-\sqrt{a}+\delta}\left[e^{2\int_0^{\tau_-} H(s)ds}; \tau_- < S \wedge \tau_+\right] \\ &\geq I(a)e^{-2\sqrt{a}S}\mathbf{P}_{-\sqrt{a}+\delta}[\tau_- < S \wedge \tau_+]. \end{aligned}$$

Now observe that the process  $Z(t) = H(t) + \sqrt{a}$  solves

$$dZ(t) = Z(t)(Z(t) - 2\sqrt{a})dt + dB(t), \quad Z(0) = \delta,$$

so that  $Z(t) \leq U(t)$  for all  $t \in [0, \tau_-(H) \wedge \tau_+(H)]$  where  $U$  is the Ornstein-Uhlenbeck process defined by

$$dU(t) = -2(\sqrt{a} - \delta)U(t)dt + dB(t), \quad U(0) = \delta.$$

Consequently, writing  $\mathbf{P}^U$  for the law of  $U$ , we get

$$\begin{aligned} \mathbf{P}_{-\sqrt{a}+\delta}[\tau_- < S \wedge \tau_+] &\geq \mathbf{P}_\delta^U[\tau_0 < S \wedge \tau_{2\delta}] \\ &\geq \mathbf{P}_\delta^U[\tau_0 < S] - \mathbf{P}_\delta^U[\tau_0 > \tau_{2\delta}]. \end{aligned}$$

By the formulae [7, II.7.2.0.2 and II.7.2.2.2, p.542] we get the bounds

$$\mathbf{P}_\delta^U[\tau_0 < S] = 1 - O(\delta a^{\frac{1}{4}} e^{-2(\sqrt{a}-\delta)S}),$$

and

$$\mathbf{P}_\delta^U[\tau_0 > \tau_{2\delta}] = \frac{\int_0^{\sqrt{2\delta}\sqrt{\sqrt{a}-\delta}} e^{v^2} dv}{\int_0^{2\sqrt{2\delta}\sqrt{\sqrt{a}-\delta}} e^{v^2} dv} \lesssim e^{-\rho a^{1/2} \delta^2},$$

for some  $\rho > 0$ . Thus we get the asserted lower bound for the first term.

We turn to the second term. By Girsanov’s Theorem, we get

$$\begin{aligned} &\mathbb{P}_{-\sqrt{a}+\delta}[S \wedge \tau_+ \leq \tau_- < S' \wedge \tau_{++}] \\ &= I(a)\mathbf{E}_{-\sqrt{a}+\delta}\left[e^{2\int_0^{\tau_-} H(s)ds}; S \wedge \tau_+ \leq \tau_- < S' \wedge \tau_{++}\right] \\ &\leq I(a)\mathbf{E}_{-\sqrt{a}+\delta}\left[e^{2\int_0^{\tau_-} H(s)ds}; S \wedge \tau_+ \leq \tau_- < \tau_{++}\right]. \end{aligned}$$

We bound the expectation at the last line by the sum of

$$A := \mathbf{E}_{-\sqrt{a}+\delta}\left[e^{2\int_0^{\tau_-} H(s)ds}; S \leq \tau_- < \tau_+\right],$$

and

$$B := \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_+ \leq \tau_- < \tau_{++} \right],$$

and we bound these two terms separately. We start with  $A$ :

$$\begin{aligned} A &\leq e^{2(-\sqrt{a}+2\delta)S} \mathbf{P}_{-\sqrt{a}+\delta} [S \leq \tau_- < \tau_+] \\ &\leq e^{2(-\sqrt{a}+2\delta)S} \left( 1 - \mathbf{P}_{-\sqrt{a}+\delta} [\tau_- < S \wedge \tau_+] \right). \end{aligned}$$

A previous calculation showed that

$$1 - \mathbf{P}_{-\sqrt{a}+\delta} [\tau_- < S \wedge \tau_+] = O(\delta a^{\frac{1}{4}} e^{-2(\sqrt{a}-\delta)S}) + O(e^{-\rho a^{1/2} \delta^2}) = \mathcal{O}((\ln a)^{2-2C}),$$

so that we get the required bound for  $A$ . Regarding  $B$ , the proof is slightly more involved. Let us first introduce:

$$\begin{aligned} B_+ &= \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_+} H(s) ds} ; \tau_+ < \tau_- \right], \\ B_- &= \mathbf{E}_{-\sqrt{a}+2\delta} \left[ e^{2 \int_0^{\tau_- - \sqrt{a} + \delta} H(s) ds} ; \tau_- - \sqrt{a} + \delta < \tau_{++} \right]. \end{aligned}$$

Applying the strong Markov property at time  $\tau_+$  we get

$$\begin{aligned} B &= \mathbf{E}_{-\sqrt{a}+\delta} \left[ \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_+ \leq \tau_- < \tau_{++} \mid \mathcal{F}_{\tau_+} \right] \right] \\ &= \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_+} H(s) ds} ; \tau_+ \leq \tau_- ; \mathbf{E}_{-\sqrt{a}+2\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_- < \tau_{++} \right] \right] \\ &= B_+ \cdot \mathbf{E}_{-\sqrt{a}+2\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_- < \tau_{++} \right]. \end{aligned}$$

Applying the strong Markov property at the first hitting time of  $-\sqrt{a} + \delta$ , we then obtain:

$$\begin{aligned} B &= B_+ B_- \cdot \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_- < \tau_{++} \right] \\ &= B_+ B_- \cdot (B + D), \end{aligned}$$

where

$$D = \mathbf{E}_{-\sqrt{a}+\delta} \left[ e^{2 \int_0^{\tau_-} H(s) ds} ; \tau_- < \tau_+ \right] \leq 1.$$

As a consequence, we find

$$B = \frac{B_+ B_-}{1 - B_+ B_-} D \leq \frac{B_+ B_-}{1 - B_+ B_-}.$$

We claim that

$$B_+ B_- \lesssim e^{2t_a \sqrt{a} - \kappa \sqrt{a} \delta^2} \lesssim a^2 e^{-\rho \ln^4 a}. \tag{41}$$

With this claim at hand, we deduce that  $B$  is bounded by a negligible term compared to  $A$ , thus concluding the proof. We are left with the proof of the claim. We have

$$B_+ \leq \mathbf{P}_{-\sqrt{a}+\delta}[\tau_+ < \tau_-] = \mathbf{P}_\delta^U[\tau_{2\delta} < \tau_0] \lesssim e^{-\rho \sqrt{a} \delta^2}.$$

To bound  $B_-$ , we argue as follows. We set

$$E := \sup_{x \in (-\sqrt{a}+\delta, \sqrt{a}-\delta/2)} \mathbf{E}_x \left[ e^{2 \int_0^{\tau_{-\sqrt{a}+\delta}} H(s) ds} ; \tau_{-\sqrt{a}+\delta} < \tau_{++} \right].$$

By considering the two complementary events  $\tau_{-\sqrt{a}+\delta} < t_a$  and  $t_a \leq \tau_{-\sqrt{a}+\delta}$ , and by applying the Markov property at time  $t_a$  in the second case, we get

$$E \leq e^{2t_a \sqrt{a}} + E e^{2t_a \sqrt{a}} \sup_{x \in (-\sqrt{a}+\delta, \sqrt{a}-\delta/2)} \mathbf{P}_x[t_a < \tau_{-\sqrt{a}+\delta} < \tau_{++}].$$

The proof of Lemma 5.1 ensures that, if  $H$  starts from  $\sqrt{a} - \delta/2$ , then we have  $\tau_{-\sqrt{a}+\delta/2} < t_a$  on the event  $\{\sup_{[0,t_a]} |B(t)| \leq c_0 \ln a/a^{1/4}\}$ , and therefore  $\tau_{-\sqrt{a}+\delta} < t_a$  on the same event. By monotonicity, this remains true if  $H$  starts from any point in  $(-\sqrt{a} + \delta, \sqrt{a} - \delta/2)$ . Recall that  $\mathbf{P}[\sup_{[0,t_a]} |B(t)| > c_0 \ln a/a^{1/4}] \lesssim a^{-c_0^2/2}$ . Consequently, choosing  $c_0$  large enough we get the crude bound

$$B_- \leq E \leq \frac{e^{2t_a \sqrt{a}}}{1 - e^{2t_a \sqrt{a}} \mathbf{P}[\sup_{[0,t_a]} |B(t)| > c_0 \ln a/a^{1/4}]} \lesssim a^2,$$

uniformly over all  $a$  large enough which concludes the proof of the claim. □

**Proof of Proposition 3.3-(i) and (ii)** Statement (i) is a consequence of Lemma 4.7. Regarding (ii), we argue as follows.

After its first hitting time of  $\sqrt{a} - \delta$ , we decompose the path  $X_a$  into two types of “bridges”:

- those that start at  $\sqrt{a} - \delta$ , hit  $\sqrt{a}$  before  $-\sqrt{a}$  and are stopped at their next hitting time of  $\sqrt{a} - \delta$ ,
- those that start at  $\sqrt{a} - \delta$ , hit  $-\sqrt{a}$  before  $\sqrt{a}$ , and are stopped at their next hitting time of  $\sqrt{a} - \delta$  (possibly after an explosion).

We are interested in the first bridge hitting  $-\sqrt{a}$ . It follows the conditional law  $\mathbb{P}_{\sqrt{a}-\delta}(\cdot \mid \tau_{-\sqrt{a}} < \tau_{\sqrt{a}})$  up to its ending time. We first show that it does not hit  $\sqrt{a} - \delta/2$  before its first hitting time of  $-\sqrt{a}$  with large probability. Indeed if we bound  $\mathbb{P}_{\sqrt{a}-\delta}[\tau_{\sqrt{a}-\delta/2} < \tau_{-\sqrt{a}}]$  by 1, then we get

$$\mathbb{P}_{\sqrt{a}-\delta}[\tau_{-\sqrt{a}} < \tau_{\sqrt{a}-\delta/2} \mid \tau_{-\sqrt{a}} < \tau_{\sqrt{a}}] \geq 1 - \frac{\mathbb{P}_{\sqrt{a}-\delta/2}[\tau_{-\sqrt{a}} < \tau_{\sqrt{a}}]}{\mathbb{P}_{\sqrt{a}-\delta}[\tau_{-\sqrt{a}} < \tau_{\sqrt{a}}]}.$$



Using the scale function (29) and the fact that there exists  $c > 0$  such that  $V(y + \delta/2) < V(y) - c\sqrt{a}\delta^2$  uniformly over all  $y \in [\sqrt{a} - \delta, \sqrt{a} - \delta/2]$ , we deduce that

$$\mathbb{P}_{\sqrt{a}-\delta}[\tau_{-\sqrt{a}} < \tau_{\sqrt{a}-\delta/2} \mid \tau_{-\sqrt{a}} < \tau_{\sqrt{a}}] \geq 1 - \exp(-c(\ln a)^4).$$

Therefore, combining Lemmas 5.1, 5.2, 5.3 and this last inequality, we obtain that there exists  $C > 0$  such that with probability larger than  $1 - 1/\ln a$ , the trajectory of the first bridge hitting  $-\sqrt{a}$  stays below  $\sqrt{a} - \delta/2$  and:

- satisfies the complement of the event  $E(C)$ ,
- hits 0 before time  $(3/8)t_a + C \ln \ln a/\sqrt{a}$  but after time  $(3/8)t_a - C \ln \ln a/\sqrt{a}$ , and hits  $-\sqrt{a} + \delta$  before time  $(3/4)t_a + 2C \ln \ln a/\sqrt{a}$  but after time  $(3/4)t_a - 2C \ln \ln a/\sqrt{a}$ ,
- stays below  $-\sqrt{a} + 2\delta$  on the time interval  $[\tau_{-\sqrt{a}+\delta}, \tau_{-\sqrt{a}}]$  and  $\tau_{-\sqrt{a}} - \tau_{-\sqrt{a}+\delta} < C \ln \ln a/\sqrt{a}$ .

Therefore, for  $X_a$  in the time interval  $[t_a, \theta_a]$ , we deduce that:

- The first portion of the trajectory from  $t_a$  until its first hitting time of  $\sqrt{a} - \delta$  stays in the interval  $[\sqrt{a} - \delta, \sqrt{a}]$  while the function  $t \mapsto -\sqrt{a} \tanh(-\sqrt{a}(t - \nu_a))$  is in  $[\sqrt{a} - (\ln a)^{4C}/a^{1/4}, \sqrt{a}]$  for  $t \leq \nu_a - (3/8)t_a + C \ln \ln a/\sqrt{a}$ . Since  $\nu_a \geq (3/8)t_a - C \ln \ln a/\sqrt{a}$ , we deduce that these two curves are at distance smaller than  $C\sqrt{a}/\ln a$  of each other (in fact, they are much closer).
- The next portion after its first hitting time of  $\sqrt{a} - \delta$  stays below  $\sqrt{a} - \delta/2$  and takes a time  $3/8 t_a + O(\ln \ln a/\sqrt{a})$  to reach 0 and then a time of the same order to reach  $-\sqrt{a} + \delta$ , and stays at a distance of order at most  $\sqrt{a}/\ln a$  from  $\sqrt{a} \tanh(-\sqrt{a}(t - \tau_0))$  where  $\tau_0$  is the first hitting time of 0 after  $t_a$ . This implies that the first and the last hitting times of 0 of this portion of the trajectory are very close to each other (at distance of order at most  $1/(\sqrt{a} \ln a)$ ): hence, we can replace  $\tau_0$  by  $\nu_a$  in the hyperbolic tangent without modifying the order of magnitude of the bound.
- The last portion of the trajectory stays below  $-\sqrt{a} + 2\delta$  and takes at most  $C \ln \ln a/\sqrt{a}$  to hit  $-\sqrt{a}$ .

This yields (ii). □

### 5.2 Coupling of $X_a$ and $X_{a+\varepsilon}$

The next lemma controls the difference between  $X_a$  and  $X_{a+\varepsilon}$ , and yields in particular the proof of Proposition 3.5.

**Lemma 5.4** (Difference between  $X_a$  and  $X_{a+\varepsilon}$ ) *Take  $\varepsilon \in (0, 1]$ . There exists  $C > 0$ , such that with probability at least  $1 - O(1/\ln a)$ , the following holds true:*

$$\begin{aligned} \sup_{t \in [\nu_a - (1/16)t_a, \nu_a + (1/16)t_a]} |X_a(t) - X_{a+\varepsilon}(t)| &< 1, \\ \sup_{t \in [\nu_a + (1/16)t_a, \theta_a - (1/16)t_a]} X_{a+\varepsilon}(t) &< -\sqrt{a} + C a^{3/7}, \end{aligned}$$

$$\sup_{t \in [\theta_a - (1/16)t_a, \theta_a]} X_{a+\varepsilon}(t) \leq \sqrt{a} - 1.$$

**Remark 5.5** Those bounds are not optimal because we overestimate the difference between  $X_a$  and  $X_{a+\varepsilon}$  at the initial time where  $X_a$  starts its descent to  $-\sqrt{a}$ . We indeed expect that when  $\varepsilon$  is of order  $1/\sqrt{a}$ , the difference between  $X_a$  and  $X_{a+\varepsilon}$  is of order  $1/a^{1/4}$  around time  $\theta_a$ .

**Proof** Recall the decomposition of the process  $X_a$  from the previous proof and let  $\sigma_a$  be the starting time of the bridge that starts at  $\sqrt{a} - \delta$  and hits  $-\sqrt{a}$  before  $\sqrt{a}$ . We introduce the process  $Z(t) := X_{a+\varepsilon}(\sigma_a + t) - X_a(\sigma_a + t)$ , which solves

$$dZ(t) = -Z(t)(X_a(\sigma_a + t) + X_{a+\varepsilon}(\sigma_a + t))dt + \varepsilon dt.$$

Since  $X_a(\sigma_a + t) \leq X_{a+\varepsilon}(\sigma_a + t)$  until  $X_a$  explodes, we deduce that

$$dZ(t) \leq -2X_a(\sigma_a + t)Z(t)dt + \varepsilon dt,$$

or, written in its integrated form

$$Z(t) \leq Z(0)e^{-2 \int_0^t X_a(\sigma_a+s)ds} + \varepsilon \int_0^t e^{-2 \int_s^t X_a(\sigma_a+r)dr} ds.$$

By Lemma 5.1, we know that with probability at least  $1 - O(a^{-c})$ , the process  $X_a(\sigma_a + t)$  is bounded from below by  $Z_-(t) - M$  until it hits  $-\sqrt{a} + \delta$ , where  $Z_-$  was introduced in the proof of Lemma 5.1. Since  $e^{2Mt} \leq 2$  for all  $t \in [0, t_a]$  and all  $a$  large enough, we have the bound

$$Z(t) \leq 2Z(0)e^{-2 \int_0^t Z_-(s)ds} + 2\varepsilon \int_0^t e^{-2 \int_s^t Z_-(r)dr} ds.$$

A simple integration yields for all  $t \geq 0$

$$\int_0^t Z_-(s)ds = \frac{1}{(1 - \kappa)^2} \ln \frac{\cosh \sqrt{a_-} A_-}{\cosh \sqrt{a_-}(t - A_-)}.$$

where  $a_- := a(1 + \kappa)(1 - \kappa)^2$ .

Similarly, we get

$$\begin{aligned} \int_0^t e^{-2 \int_s^t Z_-(r)dr} ds &= (\cosh \sqrt{a_-}(t - A_-))^{2/(1-\kappa)^2} \int_0^t \frac{1}{(\cosh \sqrt{a_-}(s - A_-))^{2/(1-\kappa)^2}} ds \\ &\leq (\cosh \sqrt{a_-}(t - A_-))^{2/(1-\kappa)^2} \int_0^t \frac{1}{(\cosh \sqrt{a_-}(s - A_-))^{2/(1-\kappa)^2}} ds \\ &\leq \frac{1}{\sqrt{a_-}} (\cosh(\sqrt{a_-}(t - A_-)))^{2/(1-\kappa)^2} \left( \tanh \sqrt{a_-}(t - A_-) + \tanh \sqrt{a_-} A_- \right). \end{aligned}$$

Fix  $\rho \in (0, 3/4)$ . For all  $t \in [\rho t_a, (\frac{3}{4} - \rho)t_a]$  we have

$$e^{-2 \int_0^t Z_-(s) ds} = \left( \frac{\cosh \sqrt{a_-}(t - A_-)}{\cosh \sqrt{a_-} A_-} \right)^{\frac{2}{(1-\kappa)^2}} \lesssim e^{-2\sqrt{a} \frac{(1+\kappa)^{1/2}}{1-\kappa} (t \wedge (2A_- - t))} \leq a^{-2\rho + o(1)}. \tag{42}$$

Additionally, we have for all  $t \in [\rho t_a, (\frac{3}{4} - \rho)t_a]$

$$\varepsilon \int_0^t e^{-2 \int_s^t Z_-(r) dr} ds \leq \varepsilon \frac{2}{\sqrt{a_-}} e^{\frac{2}{(1-\kappa)^2} \sqrt{a_-} |t - A_-|} \leq \varepsilon a^{\frac{1}{4} - 2\rho + o(1)}. \tag{43}$$

From (42) and (43), we deduce that

$$\forall t \in \left[ \rho t_a, \left( \frac{3}{4} - \rho \right) t_a \right], \quad Z(t) \leq 2Z(0)a^{-2\rho + o(1)} + \varepsilon a^{\frac{1}{4} - 2\rho + o(1)}.$$

By Lemmas 5.1, 5.2 and 5.3, we know that  $v_a - \sigma_a$  and  $\theta_a - v_a$  are of order  $3/8t_a + O(\ln \ln a / \sqrt{a})$  with probability  $1 - O(1/\ln a)$ . Furthermore,  $X_{a+\varepsilon}(\sigma_a) < 10\sqrt{a}$  with a huge probability so that  $Z(0) < 10\sqrt{a}$ . We thus easily get the first bound of the statement of the lemma.

Notice that we also deduce that there exists  $C$  such that for all  $a$  large enough

$$\sup_{t \in [v_a + (1/16)t_a, \theta_a - (1/16)t_a]} X_{a+\varepsilon}(t) < -\sqrt{a} + C a^{3/7},$$

which gives the second bound.

We turn to the third bound, for which we control  $X_{a+\varepsilon}$  on the time interval  $[\theta_a - (1/16)t_a, \theta_a]$ . We have:

$$\begin{aligned} dZ(t) &= -2X_a(t)Z(t)dt - Z(t)^2 dt + \varepsilon dt \leq 2\sqrt{a}Z(t)dt - Z(t)^2 dt + \varepsilon dt \\ &= \left( a + \varepsilon - (Z(t) - \sqrt{a})^2 \right) dt, \end{aligned}$$

for all  $t \leq \theta_a$ . Notice that  $X_a = -\sqrt{a} + o(1)$  on this time interval. Consequently, for all  $t \in [\theta_a - \frac{1}{16}t_a, \theta_a]$  we have the trivial bound:

$$X_{a+\varepsilon}(t) \leq \sqrt{a + \varepsilon} \tanh(\sqrt{a + \varepsilon}(t - A_\varepsilon)) + o(1),$$

where  $A_\varepsilon$  is such that the r.h.s. coincides with  $X_{a+\varepsilon}(\theta_a - \frac{1}{16}t_a)$  at time  $\theta_a - \frac{1}{16}t_a$ . As  $X_{a+\varepsilon}(\theta_a - \frac{1}{16}t_a) = -\sqrt{a} + O(a^{3/7})$ , we deduce the third bound of the statement.  $\square$

**Lemma 5.6** *Take  $\varepsilon \in (a^{-2/3}, 1)$ . Assume that  $X_a(t)$  and  $X_{a+\varepsilon}(t)$  lie in  $[\sqrt{a}/2, 3\sqrt{a}/2]$  for all  $t \in [0, 2t_a]$ . If  $X_{a+\varepsilon}(0) < X_a(0)$ , then by time  $2t_a$  and for all  $a$  large enough,  $X_{a+\varepsilon}$  passes above  $X_a$ .*

**Proof** Assume that  $X_{a+\varepsilon}$  is below  $X_a$  up to time  $2t_a$ , then the difference  $Z(t) = X_a(t) - X_{a+\varepsilon}(t)$  is positive and satisfies for all  $t \in [0, 2t_a]$ :

$$\begin{aligned} Z(t) &= Z(0) \exp\left(-\int_0^t (X_a + X_{a+\varepsilon})(s) ds\right) - \varepsilon \int_0^t \exp\left(-\int_v^t (X_a + X_{a+\varepsilon})(s) ds\right) dv \\ &\leq Z(0) \exp(-\sqrt{a}t) - \varepsilon \int_0^t \exp(-3\sqrt{a}(t-v)) dv \\ &\leq Z(0) \exp(-\sqrt{a}t) - \frac{\varepsilon}{3\sqrt{a}}(1 - e^{-3\sqrt{a}t}). \end{aligned}$$

By assumption  $Z(0) \leq \sqrt{a}$ . At time  $t = 2t_a$ , we thus find

$$Z(t) \leq a^{-\frac{3}{2}} - \frac{\varepsilon}{3\sqrt{a}}(1 - o(1)).$$

We get  $Z(t) < 0$  for all  $a$  large enough, thus raising a contradiction. □

### 5.3 Bounds after time $\theta_a$

**Lemma 5.7** *As  $a \rightarrow +\infty$ , we have  $\mathbb{P}_{-\sqrt{a}}(\tau_{-\sqrt{a}-\delta} < \tau_{-\sqrt{a}+\delta}) \rightarrow 1/2$ . Furthermore, there exists  $C > 0$  such that for all  $a$  large enough and all  $x \in [-\sqrt{a} - \delta, -\sqrt{a} + \delta]$ ,*

$$\mathbb{E}_x[\tau_{-\sqrt{a}+\delta} \wedge \tau_{-\sqrt{a}-\delta}] \leq C \frac{\ln \ln a}{\sqrt{a}}. \tag{44}$$

Consequently,

$$\begin{aligned} \mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}-\delta} > \frac{(\ln \ln a)^2}{4\sqrt{a}} \mid \tau_{-\sqrt{a}-\delta} < \tau_{-\sqrt{a}+\delta}] &\leq \frac{12C}{\ln \ln a}, \\ \text{and } \mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}+\delta} > \frac{(\ln \ln a)^2}{4\sqrt{a}} \mid \tau_{-\sqrt{a}+\delta} < \tau_{-\sqrt{a}-\delta}] &\leq \frac{12C}{\ln \ln a}. \end{aligned}$$

**Proof** Recall the identity (29). Since  $V(y) = V(-\sqrt{a}) - (y + \sqrt{a})^2 \sqrt{a} + (y + \sqrt{a})^3 / 3$ , a simple computation yields the following asymptotics:

$$\begin{aligned} S(-\sqrt{a} + \delta) - S(-\sqrt{a}) &= \int_{-\sqrt{a}}^{-\sqrt{a}+\delta} \exp(2V_a(y)) dy \sim \frac{\sqrt{\pi}}{2\sqrt{2}} a^{-1/4} \exp\left(\frac{4}{3}a^{3/2}\right), \\ S(-\sqrt{a} + \delta) - S(-\sqrt{a} - \delta) &= \int_{-\sqrt{a}-\delta}^{-\sqrt{a}+\delta} \exp(2V_a(y)) dy \sim \frac{\sqrt{\pi}}{\sqrt{2}} a^{-1/4} \exp\left(\frac{4}{3}a^{3/2}\right). \end{aligned} \tag{45}$$

Therefore  $\mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}-\delta} < \tau_{-\sqrt{a}+\delta}] \rightarrow 1/2$ . We now estimate the expectation of the time it takes to exit the interval  $(-\sqrt{a} - \delta, -\sqrt{a} + \delta)$  when starting at  $x \in [-\sqrt{a} - \delta, -\sqrt{a} + \delta]$ . Applying [35, Th VII.3.6], we find:

$$\begin{aligned} \mathbb{E}_x[\tau_{-\sqrt{a}-\delta} \wedge \tau_{-\sqrt{a}+\delta}] &= 2 \mathbb{P}_x[\tau_{-\sqrt{a}-\delta} < \tau_{-\sqrt{a}+\delta}] \int_{-\sqrt{a}-\delta}^x \frac{S(y) - S(-\sqrt{a} - \delta)}{S'(y)} dy \\ &\quad + 2 \mathbb{P}_x[\tau_{-\sqrt{a}+\delta} < \tau_{-\sqrt{a}-\delta}] \int_x^{-\sqrt{a}+\delta} \frac{S(-\sqrt{a} + \delta) - S(y)}{S'(y)} dy. \end{aligned}$$

Let us bound the first term on the r.h.s., the second term can be bounded in the same way. Using the change of variable  $y = -\sqrt{a} + u/a^{1/4}$ , and neglecting the cubic terms in the potential we get:

$$\int_{-\sqrt{a}-\delta}^x \frac{S(y) - S(-\sqrt{a} - \delta)}{S'(y)} dy \leq \frac{2}{\sqrt{a}} \int_{-a^{1/4}\delta}^{(x+\sqrt{a})a^{1/4}} \exp(2u^2) \int_{-a^{1/4}\delta}^u \exp(-2v^2) dv du.$$

If  $(x + \sqrt{a})a^{1/4}$  is bounded from above, the simple inequality  $\int_y^\infty \exp(-2u^2) du \leq \exp(-2y^2)/(4y)$  that holds for  $y > 0$  allows to prove that the term is of order  $\ln(\delta a^{1/4})/\sqrt{a}$ . If  $(x + \sqrt{a})a^{1/4} \rightarrow +\infty$ , the inequality

$$\mathbb{P}_x[\tau_{-\sqrt{a}-\delta} < \tau_{-\sqrt{a}+\delta}] \leq C_0 \int_{(x+\sqrt{a})a^{1/4}}^{\delta a^{1/4}} \exp(-2y^2) dy$$

allows to get the same bound. This yields (44). Thanks to Markov’s inequality and since  $\mathbb{P}_{-\sqrt{a}}(\tau_{-\sqrt{a}\pm\delta} < \tau_{-\sqrt{a}\mp\delta}) > 1/3$  for all  $a$  large enough, we get:

$$\begin{aligned} \mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}\pm\delta} > \frac{(\ln \ln a)^2}{4\sqrt{a}} \mid \tau_{-\sqrt{a}\pm\delta} < \tau_{-\sqrt{a}\mp\delta}] &\leq \frac{\mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}-\delta} \wedge \tau_{-\sqrt{a}+\delta} > (\ln \ln a)^2/(4\sqrt{a})]}{\mathbb{P}_{-\sqrt{a}}[\tau_{-\sqrt{a}\pm\delta} < \tau_{-\sqrt{a}\mp\delta}]} \\ &\leq \frac{12C}{\ln \ln a}, \end{aligned}$$

as required. □

We need a last lemma that controls the time needed by the diffusion  $X_a$  to return to  $\sqrt{a}$  when it starts from  $-\sqrt{a} + \delta$ .

**Lemma 5.8** *There exists  $C > 0$  such that for all  $a$  large enough, we have:*

$$\mathbb{P}_{-\sqrt{a}+\delta} \left[ \tau_{\sqrt{a}} \leq \frac{3}{4}t_a + C \frac{\ln \ln a}{\sqrt{a}}, \tau_{\sqrt{a}} < \tau_{-\sqrt{a}} \right] \geq 1 - \mathcal{O}(1/\ln a).$$

**Proof** It suffices to introduce  $Z := X - B$  and adapt the proof of Lemmas 5.1, 5.2 and 5.3. □

This readily implies the last event (iii) of Proposition 3.3:

**Proof of (iii) of Proposition 3.3** After time  $\theta_a$ , the process  $X_a$  has the law of the diffusion starting from  $-\sqrt{a}$ . If it exits the interval  $[-\sqrt{a} - \delta, -\sqrt{a} + \delta]$  through  $-\sqrt{a} - \delta$  then by Lemma 5.7 it does so in a time smaller than  $(\ln \ln a)^2 / (4\sqrt{a})$  with probability at least  $1 - O(1/\ln \ln a)$ , and by Lemma 4.3 it stays below  $-\sqrt{a} + 1$  and goes to  $-\infty$  within a time  $(3/8)t_a$ , with probability at least  $1 - \exp(-(\ln \ln a)^2)$ . On the other hand, if it exits through  $-\sqrt{a} + \delta$ , then by Lemma 5.8 it reaches  $\sqrt{a}$  without hitting  $-\sqrt{a}$  within a time  $(3/4)t_a + C \ln \ln a / \sqrt{a}$  with probability at least  $1 - O(1/\ln a)$ .  $\square$

**Proof of Proposition 3.6** The r.v.  $\theta \wedge L$  is a stopping time in the filtration  $\mathcal{F}_t, t \geq 0$  of the underlying Brownian motion  $B$ . By the strong Markov property, the process  $(B(t + \theta \wedge L) - B(\theta \wedge L), t \geq 0)$  is a standard Brownian motion, independent from  $\mathcal{F}_{\theta \wedge L}$ . Hence, conditionally given  $\theta \wedge L$ , the process  $(\hat{X}_a(t), t \in [0, L - \theta \wedge L])$  has the law of the time-reversed diffusion stopped at the deterministic time  $L - \theta \wedge L$ . The two bounds of the statement follow from the same type of arguments as those presented in the proofs of Lemmas 4.5 and 4.7. In particular, we introduce a stationary diffusion  $\hat{Y}$  (with the same law as the stationary diffusion  $-Y$ ) driven by  $\hat{B}$ . We define the event

$$\mathcal{A} := \{ \hat{Y}((L - t) \vee 0) \leq -\sqrt{a} + h_a/2 \text{ for all } t \in [\theta, \theta + 11t_a] \}.$$

From the estimates on the invariant measure collected in Lemma 4.1 and using a comparison with a reflected Brownian motion (as in the proof of Proposition 2.6), we deduce that there exists  $\rho > 0$  such that  $\mathbb{P}[\mathcal{A}] > 1 - 1/\ln a$  for all  $a$  large enough. Denote by  $\mathcal{B}$  the event on which  $\hat{X}_a$  and  $\hat{Y}$  remain at a distance  $h_a/2$  from each other on the interval  $[(3/8)t_a, \tau]$  where  $\tau$  is the first hitting time of  $3\sqrt{a}$  by  $\hat{Y}$  and both of them explode after time  $\tau$  within a time  $O(1/\sqrt{a})$ . The proof of Lemma 4.7 ensures that  $\mathbb{P}[\mathcal{B}] \geq 1 - 1/\ln a$ .

Define now the event  $\mathcal{C}$  on which

$$\int_{\theta \wedge L}^t \hat{Y}(L - s) ds \leq -\sqrt{a} + h_a/2, \quad \forall t \in [\theta \wedge L, L].$$

We claim that there exists  $c > 0$  such that  $\mathbb{P}[\mathcal{C} | \mathcal{F}_{\theta \wedge L}] > 1 - \exp(-c(\ln a)^2)$ , and consequently  $\mathbb{P}[\mathcal{C}] > 1 - \exp(-c(\ln a)^2)$ . Indeed, if we introduce the events

$$E_k := \left\{ \int_{\theta \wedge L}^{\theta \wedge L + 2^{-k}(L - \theta \wedge L)} \mathbf{1}_{\{-\hat{Y}(L-s) \notin I_a(1/4)\}} ds \geq \frac{1}{a} \right\}, \quad k \geq 1,$$

then, by the stationarity of  $\hat{Y}$ , we deduce that  $\mathbb{P}[E_k | \mathcal{F}_{\theta \wedge L}] \leq a e^{-(\ln a)^2/16}$  so that the proof of Lemma 4.5 carries through *mutatis mutandis* and yields the asserted bound on the conditional probability of  $\mathcal{C}$ .

Let  $\mathcal{D}$  be the event on which  $\hat{Y}$  explodes to  $+\infty$  within a time of order  $1/\sqrt{a}$  once it has hit  $3\sqrt{a}$ . By Lemma 4.3,  $\mathbb{P}(\mathcal{D}) > 1 - \exp(-(\ln \ln a)^2)$ .

Therefore, on the event  $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C} \cap \mathcal{D}$ , if  $\hat{\zeta}_a(1) > L - \theta - 10t_a$  and  $\theta < L$  then  $\hat{Y}$  hasn't reached  $3\sqrt{a}$  by time  $L - \theta - 10t_a$ : indeed, if it had then by  $\mathcal{D}$  it would explode before time  $L - \theta$  and this would contradict  $\mathcal{A}$ . Using  $\mathcal{A}$ , this in turn ensures that  $\tau > L - \theta$ . Since  $\hat{X}_a$  remains below  $\hat{Y}$  up to the first explosion time of the latter, the bound of event  $\mathcal{A}$  yields:

$$\hat{X}_a(L - t) \leq -\sqrt{a} + h_a, \quad \forall t \in [\theta, \theta + 10t_a],$$

so that  $\hat{X}_a$  does not explode before time  $L - \theta$ . Moreover the bound of event  $\mathcal{C}$  combined with the condition of event  $\mathcal{B}$  yields the second bound of the proposition.  $\square$

**5.4 Proof of parts (b)–(i), (b)–(ii) and (b)–(iii) of Proposition 3.9**

Let us fix  $\varepsilon > 0$ . Let  $a \in M_L$ . Recall that by Proposition 2.4, the explosion times of  $X_a$  (resp.  $X_a^j$ ), rescaled by  $m(a_L) = L$ , converge to a Poisson point process on  $\mathbf{R}_+$  (resp. on  $[t_j^n, +\infty)$ ) of intensity bounded from above by  $e^{4/\varepsilon}$  as  $a \geq a_L - 1/(\varepsilon\sqrt{a_L})$ . Consequently, with probability  $1 - O(2^{-n})$  for all  $L$  large enough,  $X_a$  does not explode on  $[(t_j^n - 2 \cdot 2^{-2n}L) \vee 0, (t_j^n + 2 \cdot 2^{-2n}L) \wedge L]$ ,  $X_a^j$  explodes at most once per interval  $[t_j^n, t_{j+3}^n]$ . This yields (b)-(i) and (b)-(iii) for all  $a \in M_L$  with probability  $1 - O(2^{-n})$  for all  $L$  large enough (since there are of order  $1/\varepsilon^2$  points in  $M_L$ ).

It is easy to generalize [3, Thm 3.3] to show that the first hitting time of  $-\sqrt{a}$ , rescaled by  $m(a)/2$ , of the diffusion  $X_a$  starting from  $\sqrt{a}$  converges to an exponential r.v. of parameter 1 when  $a \rightarrow \infty$  (its Laplace transform satisfies a similar fixed point equation as in [3, Proposition 3.3] replacing  $-\infty$  in the first integral by  $-\sqrt{a}$  and taking  $y \geq -\sqrt{a}$ ). The same proof then carries through, noting that we have to divide  $m(a)$  by 2 so that the recursive integrals  $R_n(y, a)$  converges to 1). As a consequence, we have the following counterpart of Proposition 2.4: the first hitting times of  $-\sqrt{a}$  of the successive excursions of  $X_a$ , rescaled by  $m(a)$ , converge as  $a \rightarrow \infty$  to a Poisson point process on  $\mathbf{R}_+$  of intensity 2. Consequently the first hitting times of  $-\sqrt{a}$  of the successive excursions of  $X_a$  (resp.  $X_a^j$ ), rescaled by  $m(a_L) = L$ , converge as  $L \rightarrow \infty$  to a Poisson point process on  $\mathbf{R}_+$  (resp. on  $[t_j^n, \infty)$ ) of intensity bounded from above by  $2e^{4/\varepsilon}$ . Therefore for  $j = 0$  and  $j = 2^n - 1$ ,  $X_a$  and  $X_a^j$  do not hit  $-\sqrt{a}$  on  $[t_j^n, t_{j+1}^n]$  with probability at least  $1 - O(2^{-n})$ . For any other given  $j$ ,  $X_a$  and  $X_a^j$  hit at most once  $-\sqrt{a}$  with probability at least  $1 - O(2^{-2n})$  so that, taking a union bound over all such  $j$ , we obtain a probability at least  $1 - O(2^{-n})$ . Hence (b)-(ii) is satisfied with probability at least  $1 - O(2^{-n})$  for all  $L$  large enough.

**5.5 Proof of part (b)–(iv) of Proposition 3.9**

Let  $a_\ell$  be the smallest point in  $M_L$ . In the next lemma, we show that for  $a > a_\ell$  in  $M_L$ , if  $X_a^j$  explodes then  $X_{a_\ell}^j$  explodes roughly at the same time, and this proves (b)-(iv).

**Lemma 5.9** Fix  $n \geq 1$  and take  $a \in M_L$ . The following holds for all  $j \in \{0, \dots, 2^n - 1\}$  with probability at least  $1 - O(1/\ln \ln a_L)$ . If  $X_{a_\ell}^j$  makes at most one excursion to  $-\sqrt{a_\ell}$  on  $[t_j^n, t_{j+1}^n]$ , then:

- (A) on  $[t_j^n, t_{j+1}^n]$  the number of excursions of  $X_a^j$  to  $-\sqrt{a}$  is smaller than or equal to the number of excursions of  $X_{a_\ell}^j$  to  $-\sqrt{a_\ell}$ ,
- (B) if  $X_a^j$  hits  $-\sqrt{a}$  on  $[t_j^n, t_{j+1}^n]$  and then explodes without coming back to  $\sqrt{a}$ , then so does  $X_{a_\ell}^j$  and their explosion times lie at a distance at most  $(\ln \ln a_L)^2 / (3\sqrt{a_L})$  from each other.

**Proof** Notice that the difference  $a - a_\ell$  belongs to  $(a_L^{-2/3}, 1)$ . We are going to describe the behavior of the two trajectories  $X_a^j, X_{a_\ell}^j$  by combining several estimates obtained in previous lemmas. The bounds on the probabilities obtained in these lemmas ensure that all what follows happens with probability greater than  $1 - O(1/\ln \ln a_L)$ . The diffusion  $X_a^j$  remains above  $X_{a_\ell}^j$  until the first explosion time of  $X_{a_\ell}^j$ , and cannot reach  $-\sqrt{a}$  before the first hitting time of  $-\sqrt{a_\ell}$  by  $X_{a_\ell}^j$ . Let us denote this first hitting time  $\theta_{a_\ell}^j$ . At time  $\theta_{a_\ell}^j$ , by the estimates of Proposition 3.5 the diffusion  $X_a^j$  lies in  $[-\sqrt{a_\ell}, \sqrt{a_\ell} - 1] \subset [-\sqrt{a}, \sqrt{a_\ell} - 1]$ . For convenience, let us distinguish two cases: the diffusion is either in the interval  $[-\sqrt{a}, -\sqrt{a} + \delta]$  or in the interval  $[-\sqrt{a} + \delta, \sqrt{a_\ell} - 1]$ , where  $\delta = \ln^2(a)/a^{1/4}$ .

In the first case, thanks to Lemma 5.7, it takes a time smaller than  $(\ln \ln a_L)^2 / (4\sqrt{a_L})$  to exit the interval  $[-\sqrt{a} - \delta, \sqrt{a} + \delta]$  with probability greater than  $1 - O(1/\ln \ln a)$ . If it exits through  $-\sqrt{a} - \delta$ , then it takes a time  $(3/8)t_L + O(\ln \ln a_L / \sqrt{a_L})$  to explode to  $-\infty$  (without hitting  $-\sqrt{a}$ ) by Lemma 4.3. From the ordering of the diffusions, this implies that  $X_{a_\ell}^j$  explodes as well and that their explosion times are at distance at most  $(\ln \ln a_L)^2 / (3\sqrt{a_L})$  from one another. If it exits through  $-\sqrt{a} + \delta$ , then we get to the second case.

In the second case, thanks to Lemma 5.8, it takes a time smaller than  $(3/4)t_L + O(\ln \ln a_L / \sqrt{a_L})$  for the diffusion to reach  $\sqrt{a}$  without hitting  $-\sqrt{a}$ . Moreover, after this return time, it stays in the region  $[\sqrt{a_\ell}/2, (3/2)\sqrt{a_\ell}]$  during a time greater than  $5t_L$  (notice that this happens with a huge probability). Similarly, the diffusion  $X_{a_\ell}^j$  hits  $\sqrt{a_\ell}$  (with or without explosion) before time  $\theta_a^j + (3/4)t_L + o(t_L)$  and stays in the interval  $[\sqrt{a}/2, (3/2)\sqrt{a_\ell}]$  during a time greater than  $5t_L$ . We can therefore apply Lemma 5.6 which states that  $X_a^j$  passes above  $X_{a_\ell}^j$  before time  $\theta_{a_\ell}^j + 5t_L$ , thus preventing  $X_a^j$  to make a second excursion to  $-\sqrt{a}$  before  $X_{a_\ell}^j$  makes itself a second excursion to  $-\sqrt{a_\ell}$ .

Therefore, with probability  $1 - O(1/\ln \ln a_L)$  if  $X_{a_\ell}^j$  makes only one excursion to  $-\sqrt{a_\ell}$  on  $[t_j^n, t_{j+1}^n]$  then (A) and (B) are satisfied. Since there are  $2^n$  different values  $j$ , the statement follows.  $\square$



### 6 Neumann boundary conditions

When the operator is endowed with Neumann boundary conditions, we have  $\varphi'_k(0) = \varphi'_k(L) = 0$  so that necessarily  $\varphi_k(0), \varphi_k(L) \neq 0$  (otherwise,  $\varphi_k$  would be identically zero). Consequently, the corresponding Riccati transforms  $\chi_k^{(N)}$  start and end at 0. Therefore, we need to start our diffusions from 0 as well. To avoid confusions, we let  $X_a^{(N)}$  and  $\hat{X}_a^{(N)}$  be the analogues of  $X_a$  and  $\hat{X}_a$  but starting from 0. Notice that we have the following almost sure equivalences:

- $a > -\lambda_1^{(N)}$  if and only if  $X_a^{(N)}$  does not explode on  $[0, L]$  and  $X_a^{(N)}(L) > 0$ ,
- $-\lambda_k^{(N)} \geq a > -\lambda_{k+1}^{(N)}$  if and only if either  $[X_a^{(N)}$  explodes  $k$  times on  $[0, L]$  and  $X_a^{(N)}(L) > 0]$  or  $[X_a^{(N)}$  explodes  $k - 1$  times on  $[0, L]$  and  $X_a^{(N)}(L) \leq 0]$ .

The strategy of proof is exactly the same as in the case of Dirichlet boundary conditions: we decompose the interval  $[0, L]$  into subintervals of size  $2^{-n}L$ , and, within these subintervals, we still consider the diffusions  $X_a^j$  that start from  $+\infty$  (and not from 0!) at time  $j2^{-n}L$ . The only changes in the proof consist in dealing with the boundary conditions of  $X_a^{(N)}$  and  $\hat{X}_a^{(N)}$ . One needs to show that with large probability we have:

- (1) The diffusion  $X_a^{(N)}$  reaches a neighborhood of  $\sqrt{a}$  very quickly:

$$\tau_x(X_a^{(N)}) = (3/8)t_a(1 + o(1)),$$

where  $x = \sqrt{a} - \frac{\ln a}{a^{1/4}}$  and

$$\sup_{t \in (0, \frac{3}{8}t_a]} \left| X_a^{(N)}(t) - \sqrt{a} \tanh(\sqrt{a}t) \right| \leq 1.$$

- (2) The diffusion  $X_a^{(N)}$  synchronizes with  $X_a^j$  for every  $j = 0, \dots, 2^n - 1$ : the content of Proposition 2.6 remains true upon replacing  $X_a$  by  $X_a^{(N)}$  and  $t_0$  by  $j2^{-n}L$ .
- (3) If  $X_a^{(N)}$  does not explode more than  $k$  times, then  $X_a^{(N)}(L) > 0$ .

The proofs of the two first estimates are essentially the same as those for Dirichlet boundary conditions. The proof of the third estimate is a simple consequence of the synchronization with the stationary diffusion proved in Lemma 4.7 for Dirichlet boundary conditions, which can be adapted to the Neumann case, and of the fact that, for a stationary diffusion  $Y$  the probability that  $Y(L) > \sqrt{a}/2$  is overwhelming. This being given, one considers a more restrictive event  $\mathcal{E}(n, \epsilon)$  than previously: one also imposes that for all  $a \in M_L$ ,  $X_a^{(N)}$  satisfies the above bounds and (b)–(ii), and similarly for  $\hat{X}_a^{(N)}$ . This event  $\mathcal{E}(n, \epsilon)$  has a probability of order at least  $1 - O(\epsilon)$ .

We now work on  $\mathcal{E}(n, \epsilon)$ . We know that for every  $a \in M_L$ , the sum over  $j$  of the number of explosions of  $X_a^j$  coincides with the number of explosions of  $X_a$ . Thanks to the additional properties collected above, this is also true for  $X_a^{(N)}$ . Since in addition

$X_a^{(N)}(L) > 0$ , the r.v.  $a_i$  and  $a'_i$ , which were defined in Sect. 3, satisfy  $a_i \leq -\lambda_i^{(N)} < a'_i$ . Since  $a'_i - a_i = \epsilon/\sqrt{a_L}$ , this immediately implies that  $(\lambda_i^{(N)} - \lambda_i)\sqrt{a_L}$  goes to 0 in probability. Then, the same monotonicity arguments as for the Dirichlet case allow to bound  $\chi_i^{(N)}$  using the diffusions  $X_{a_i}^{(N)}$ ,  $X_{a'_i}^{(N)}$  and their time reversals. In particular, the location of the maximum of  $\varphi_i^{(N)}$  is still very close to the point  $v_*$ , defined as the last zero of  $X_{a_i}^{j_*}$  where the time-interval  $[j_*2^{-n}L, (j_* + 1)2^{-n}L]$  corresponds to the additional explosion of  $X_{a_i}^{(N)}$ . This concludes the proof of Theorem 4.

**Acknowledgements** We thank the anonymous referees for their careful reading of the paper and their helpful remarks. LD thanks Romain Allez and Benedek Valkó for useful discussions. The work of LD is supported by the project MALIN ANR-16-CE93-0003. CL is grateful to Julien Reygner for several useful comments on a preliminary version of this paper. The work of CL is supported by the project SINGULAR ANR-16-CE40-0020-01.

## References

- Allez, R., Chouk, K.: The continuous Anderson Hamiltonian in dimension two. ArXiv e-prints (2015). [arXiv:1511.02718](https://arxiv.org/abs/1511.02718)
- Allez, R., Dumaz, L.: From sine kernel to Poisson statistics. *Electron. J. Probab.* **19**(114), 25 (2014). <https://doi.org/10.1214/EJP.v19-3742>
- Allez, R., Dumaz, L.: Tracy–Widom at high temperature. *J. Stat. Phys.* **156**(6), 1146–1183 (2014). <https://doi.org/10.1007/s10955-014-1058-z>
- Allez, R., Dumaz, L.: Random matrices in non-confining potentials. *J. Stat. Phys.* **160**(3), 681–714 (2015). <https://doi.org/10.1007/s10955-015-1258-1>
- Anderson, P.W.: Absence of diffusion in certain random lattices. *Phys. Rev.* **109**(5), 1492–1505 (1958). <https://doi.org/10.1103/PhysRev.109.1492>
- Biskup, M., König, W.: Eigenvalue order statistics for random Schrödinger operators with doubly-exponential tails. *Commun. Math. Phys.* **341**(1), 179–218 (2016). <https://doi.org/10.1007/s00220-015-2430-9>
- Borodin, A.N., Salminen, P.: *Handbook of Brownian Motion-Facts and Formulae. Probability and Its Applications*, 2nd edn. Birkhäuser Verlag, Basel (2002)
- Bloemendal, A., Virág, B.: Limits of spiked random matrices II. *Ann. Probab.* **44**(4), 2726–2769 (2016). <https://doi.org/10.1214/15-AOP1033>
- Carmona, R., Klein, A., Martinelli, F.: Anderson localization for Bernoulli and other singular potentials. *Commun. Math. Phys.* **108**(1), 41–66 (1987)
- Cambronero, S., McKean, H.P.: The ground state Eigenvalue of Hill's equation with white noise potential. *Commun. Pure Appl. Math.* **52**(10), 1277–1294 (1999). [https://doi.org/10.1002/\(SICI\)1097-0312\(199910\)52:10<1277::AID-CPA5>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-0312(199910)52:10<1277::AID-CPA5>3.0.CO;2-L)
- Cambronero, S., Rider, B., Ramírez, J.: On the shape of the ground state eigenvalue density of a random Hill's equation. *Commun. Pure Appl. Math.* **59**(7), 935–976 (2006). <https://doi.org/10.1002/cpa.20104>
- Dumaz, L., Virág, B.: The right tail exponent of the Tracy–Widom  $\beta$  distribution. *Ann. Inst. Henri Poincaré Probab. Stat.* **49**(4), 915–933 (2013). <https://doi.org/10.1214/11-AIHP475>
- Frisch, H.L., Lloyd, S.P.: Electron levels in a one-dimensional random lattice. *Phys. Rev.* **120**, 1175–1189 (1960). <https://doi.org/10.1103/PhysRev.120.1175>
- Fukushima, M., Nakao, S.: On spectra of the Schrödinger operator with a white Gaussian noise potential. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **37**(3), 267–274 (1976/77)
- Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, e6, 75 (2015). [arXiv:1210.2684](https://arxiv.org/abs/1210.2684). <https://doi.org/10.1017/fmp.2015.2>
- Goldshid, I.J., Molcanov, S.A., Pastur, L.A.: A random homogeneous Schrödinger operator has a pure point spectrum. *Funkcional. Anal. i Priložen.* **11**(1), 1–10 (1977). 96
- Gubinelli, M., Ugurcan, B., Zichhuber, I.: Semilinear evolution equations for the Anderson Hamiltonian in two and three dimensions. ArXiv e-prints (2018). [arXiv:1807.06825](https://arxiv.org/abs/1807.06825)

18. Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014). [arXiv:1303.5113](https://arxiv.org/abs/1303.5113). <https://doi.org/10.1007/s00222-014-0505-4>
19. Halperin, B.I.: Green's functions for a particle in a one-dimensional random potential. *Phys. Rev. (2)* **139**, A104–A117 (1965)
20. König, W.: *The Parabolic Anderson Model. Pathways in Mathematics. Random Walk in Random Potential.* Birkhäuser/Springer, Cham (2016)
21. Kallenberg, O.: *Foundations of Modern Probability. Probability and Its Applications (New York)*, 2nd edn. Springer, New York (2002)
22. Kirsch, W.: An invitation to random Schrödinger operators. In *Random Schrödinger operators*, vol. 25 of *Panor. Synthèses*, 1–119. Soc. Math. France, Paris, 2008. With an appendix by Frédéric Klopp
23. Killip, R., Nakano, F.: Eigenfunction statistics in the localized Anderson model. *Ann. Henri Poincaré* **8**(1), 27–36 (2007). <https://doi.org/10.1007/s00023-006-0298-0>
24. Killip, R., Stoiciu, M.: Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.* **146**(3), 361–399 (2009). <https://doi.org/10.1215/00127094-2009-001>
25. Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. *Commun. Math. Phys.* **78**(2), 201–246 (1980/81)
26. Kritchevski, E., Valkó, B., Virág, B.: The scaling limit of the critical one-dimensional random Schrödinger operator. *Commun. Math. Phys.* **314**(3), 775–806 (2012). <https://doi.org/10.1007/s00220-012-1537-5>
27. Labbé, C.: The continuous Anderson Hamiltonian in  $d \leq 3$ . *ArXiv e-prints* (2018). [arXiv:1809.03718](https://arxiv.org/abs/1809.03718)
28. McKean, H.P.: A limit law for the ground state of Hill's equation. *J. Stat. Phys.* **74**(5–6), 1227–1232 (1994). <https://doi.org/10.1007/BF02188225>
29. Minami, N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Commun. Math. Phys.* **177**(3), 709–725 (1996)
30. Molčanov, S.A.: The local structure of the spectrum of the one-dimensional Schrödinger operator. *Commun. Math. Phys.* **78**(3), 429–446 (1980/81)
31. Otto, F., Weber, H., Westdickenberg, M.G.: Invariant measure of the stochastic Allen–Cahn equation: the regime of small noise and large system size. *Electron. J. Probab.* **19**(23), 76 (2014). <https://doi.org/10.1214/EJP.v19-2813>
32. Ramírez, J.A., Rider, B.: Diffusion at the random matrix hard edge. *Commun. Math. Phys.* **288**(3), 887–906 (2009). <https://doi.org/10.1007/s00220-008-0712-1>
33. Ramírez, J.A., Rider, B., Virág, B.: Beta ensembles, stochastic airy spectrum, and a diffusion. *J. Am. Math. Soc.* **24**(4), 919–944 (2011). <https://doi.org/10.1090/S0894-0347-2011-00703-0>
34. Rifkind, B., Virág, B.: Eigenvectors of the critical 1-dimensional random Schrödinger operator. *ArXiv e-prints* (2016). [arXiv:1605.00118](https://arxiv.org/abs/1605.00118)
35. Revuz, D., Yor, M.: *Continuous martingales and Brownian motion.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, 3rd edn. Springer, Berlin (1999)
36. Texier, C.: Individual energy level distributions for one-dimensional diagonal and off-diagonal disorder. *J. Phys. A* **33**(35), 6095–6128 (2000). <https://doi.org/10.1088/0305-4470/33/35/303>
37. Valkó, B., Virág, B.: Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177**(3), 463–508 (2009). <https://doi.org/10.1007/s00222-009-0180-z>