

**Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig**

**Invariant measure of the stochastic Allen-Cahn  
equation: the regime of small noise and large  
system size**

by

*Felix Otto, Hendrik Weber, and Maria G. Westdickenberg*

Preprint no.: 1

2013





# INVARIANT MEASURE OF THE STOCHASTIC ALLEN-CAHN EQUATION: THE REGIME OF SMALL NOISE AND LARGE SYSTEM SIZE

FELIX OTTO, HENDRIK WEBER, AND MARIA G. WESTDICKENBERG

**ABSTRACT.** We study the invariant measure of the one-dimensional stochastic Allen-Cahn equation for a small noise strength and a large but finite system. We endow the system with inhomogeneous Dirichlet boundary conditions that enforce at least one transition from  $-1$  to  $1$ . (Our methods can be applied to other boundary conditions as well.) We are interested in the competition between the “energy” that should be minimized due to the small noise strength and the “entropy” that is induced by the large system size.

Our methods handle system sizes that are exponential with respect to the inverse noise strength, up to the “critical” exponential size predicted by the heuristics. We capture the competition between energy and entropy through upper and lower bounds on the probability of extra transitions between  $\pm 1$ . These bounds are sharp on the exponential scale and imply in particular that the probability of having one and only one transition from  $-1$  to  $+1$  is exponentially close to one. In addition, we show that the position of the transition layer is uniformly distributed over the system on scales larger than the logarithm of the inverse noise strength.

Our arguments rely on local large deviation bounds, the strong Markov property, the symmetry of the potential, and measure-preserving reflections.

## 1. INTRODUCTION

In this paper we study the unique invariant measure of the stochastically perturbed Allen-Cahn equation

$$\partial_t u_\varepsilon(t, x) = \partial_x^2 u_\varepsilon(t, x) - V'(u_\varepsilon(t, x)) + \sqrt{2\varepsilon} \eta(t, x), \quad (1.1)$$

where  $u_\varepsilon$  is a one-dimensional order parameter defined for all non-negative times  $t \in \mathbb{R}_+$  and  $x \in (-L_\varepsilon, L_\varepsilon)$ . Here  $\eta$  is a formal expression denoting space-time white noise and  $V$  is a symmetric double-well potential. The canonical choice for  $V$  is

$$V(u) = \frac{1}{4}(1 - u^2)^2,$$

although more general choices are possible (see Assumption 1.1 below). We are interested in the properties of the invariant measure for large system sizes,

$$L_\varepsilon \gg 1.$$

It is well-known that for  $\varepsilon \downarrow 0$  and fixed system size  $L$ , the invariant measure of the Allen-Cahn equation concentrates on minimizers of the energy

$$\int_{-L}^L \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx.$$

This follows from large deviation theory. In fact, even for system sizes  $L_\varepsilon$  that grow with  $\varepsilon$ , the same is true. Indeed, in [Web10] the second author proved this fact for  $L_\varepsilon \sim \varepsilon^{-\alpha}$  for any  $\alpha < \frac{2}{3}$ .

Our main goal in the current paper is to go up to interval sizes that are *exponential* with respect to  $\varepsilon^{-1}$  and, specifically, to understand the *competition between energy and entropy* that emerges in this regime. Let us first consider the effect of *energy* on the measure. The intuition is that the invariant measure can be viewed as a Gibbs measure with the given energy, i.e., that it is in some heuristic sense proportional to

$$\exp \left( -\frac{1}{\varepsilon} \int_{-L_\varepsilon}^{L_\varepsilon} \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx \right).$$

The heuristic picture then says that, because of the potential term in the energy, functions  $u$  supported on this measure are most likely to be close to one or the other minimum of  $V$  on most of  $[-L_\varepsilon, L_\varepsilon]$ . On the other hand, because of the gradient term in the energy, there is an energetic “cost”  $c_0$  for each transition between these two preferred states, making such transitions unlikely. When the system size is order-one, this is the end of the story.

Now let us consider the competing effect of *entropy* on the measure when the system size is large. Namely, the probability of finding a transition between the minima of  $V$  is increased by the fact that it is possible for the transition to occur anywhere in the system. Hence, the folklore is that the probability of finding  $n$  transition layers scales like

$$(L_\varepsilon)^n \exp \left( -\frac{nc_0}{\varepsilon} \right). \quad (1.2)$$

This competition between entropy and energy is captured (on the exponential level) in our first theorem, Theorem 1.5 below. Our second result, Theorem 1.9, then shows the uniform distribution of transitions within the domain.

As far as our methods, the central idea is that one can decompose the measure into conditional measures and the corresponding marginals in order to reduce to order-one intervals on which one can apply large deviation theory. Along the way, it is important for us to use measure-preserving *reflection arguments* that allow us to transform the underlying Brownian paths. The detailed structure of the (deterministic) energy functional is also critical in our proofs.

We will state our results in detail in Subsection 1.2 after first explaining our set-up and notation.

**1.1. Set-up and notation.** For the potential  $V$  in (1.1), we need a symmetric double-well potential with at least superlinear growth at infinity. For simplicity, we assume that the two minima of  $V$  are normalized to be at  $\pm 1$  and that the minimum value of the potential is zero. To be precise, our assumptions are:

**Assumption 1.1.**  $V$  is a smooth, even potential such that, on  $(0, \infty)$ ,  $V$  satisfies

$$\begin{aligned} V(u) &\geq 0 \quad \text{and} \quad V(u) = 0 \quad \text{iff} \quad u = 1, \\ V'(u) &= 0 \quad \text{if and only if} \quad u = 1, \\ V''(1) &> 0, \\ V(u) &\geq u^{1+\beta}/C \quad \text{for} \quad u \geq C \quad \text{for some } C < \infty \text{ and } \beta > 0. \end{aligned} \quad (1.3)$$

*Remark 1.2.* If we assume superquadratic growth on  $V$  at infinity (recall that we have *quartic* growth of the standard double well potential  $V(u) = (1 - u^2)^2/4$ ), some of our technical lemmas simplify slightly. In particular, one can remove the dependence of the minimal system size  $\ell_*$  on  $M$  in Lemmas 2.3 and 2.5.

Because of the normalization of our potential, the transitions that we are interested in are transitions between  $\pm 1$ . We make the notion precise in the following definition.

**Definition 1.3** (Up/down transition layers). *We say that  $u$  has an up transition layer on  $(x_-, x_+)$  if*

$$u(x_{\pm}) = \pm 1 \quad \text{and} \quad |u(x)| < 1 \quad \text{for all } x \in (x_-, x_+).$$

*We say that  $u$  has a down transition layer on  $(x_-, x_+)$  if the same condition holds with signs reversed, and that  $u$  has a transition layer if it has an up or down transition layer.*

For the boundary conditions on our PDE, we will work with the popular inhomogeneous Dirichlet boundary conditions

$$u_{\varepsilon}(t, \pm L_{\varepsilon}) = \pm \frac{1}{\varepsilon} \int_{-\ell_0}^{\ell_0} V(\varepsilon^{1/2}(\hat{u} - 1) + 1) dx. \quad (1.4)$$

Because of the boundary conditions, there is necessarily *one* up transition layer, and the question is whether there are additional transition layers. Moreover, if there are additional layers, they come as a pair of an up layer and a down layer. Note that our methods can also handle other boundary conditions, for instance periodic boundary conditions or Dirichlet boundary conditions that do not force a transition layer to be present.

We will denote the invariant measure of (1.1) subject to the boundary conditions (1.4) by  $\mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$  and the corresponding expectation by  $\mathbb{E}_{(-L_{\varepsilon}, L_{\varepsilon})}^{\mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}}(\cdot)$ . We will often use the fact that the measure  $\mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$  can be written as a Gaussian measure with density [Zab89]. Namely, one can express the expectation of any test function  $\Phi$  as

$$\mathbb{E}_{(-L_{\varepsilon}, L_{\varepsilon})}^{\mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}}(\Phi) = \frac{\mathbb{E}_{(-L_{\varepsilon}, L_{\varepsilon})}^{\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}} \left[ \Phi(u) \exp \left( -\frac{1}{\varepsilon} \int_{-L_{\varepsilon}}^{L_{\varepsilon}} V(u) dx \right) \right]}{\mathbb{E}_{(-L_{\varepsilon}, L_{\varepsilon})}^{\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}} \left[ \exp \left( -\frac{1}{\varepsilon} \int_{-L_{\varepsilon}}^{L_{\varepsilon}} V(u) dx \right) \right]}. \quad (1.5)$$

Here  $\mathbb{E}_{(-L_{\varepsilon}, L_{\varepsilon})}^{\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}}$  denotes the expectation with respect to the measure  $\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$ , which is the distribution of a Brownian bridge on  $(-L_{\varepsilon}, L_{\varepsilon})$  from  $-1$  to  $+1$  with variance proportional to  $\varepsilon$ . Properties of  $\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$  will be discussed in detail in Section 3.

The deterministic Allen-Cahn equation (set  $\eta = 0$  in (1.1)) is the  $L^2$ -gradient flow of the energy

$$E(u) := \int_{-L_{\varepsilon}}^{L_{\varepsilon}} \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx. \quad (1.6)$$

When we need to refer to the energy on all of  $\mathbb{R}$  or the localized energy on a subinterval, we will denote this with a subscript:

$$\begin{aligned} E_{(-\infty, \infty)}(u) &= \int_{-\infty}^{\infty} \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx, \\ E_{(-\ell, \ell)}(u) &= \int_{-\ell}^{\ell} \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx. \end{aligned} \quad (1.7)$$

As mentioned above, the energy functional will be important for understanding the invariant measure of the stochastic equation. In particular, the probability of finding transition layers will depend on the energetic “cost” of a transition layer on  $\mathbb{R}$ , that is:

$$c_0 := \inf \{ E_{(-\infty, \infty)}(u) : u(\pm\infty) = \pm 1 \}. \quad (1.8)$$

It is well known [MM77] that this cost can be computed explicitly as

$$c_0 = \int_{-1}^1 \sqrt{2V(u)} du \stackrel{\text{Ass. 1.1}}{=} 2 \int_0^1 \sqrt{2V(u)} du; \quad (1.9)$$

see the beginning of Section 2 for an explanation.

We will often refer to scaling regimes in our results. To this end, we define the following notation.

**Notation 1.4.** *The well-established theory of large deviations applies on intervals whose length is order-one with respect to  $\varepsilon$ . A main point of this paper, however, is to obtain estimates on intervals that are exponentially large with respect to  $\varepsilon$  and for which, consequently, the established theory does not apply. We therefore use a subscript of  $\varepsilon$  in order to distinguish interval lengths that are large with respect to  $\varepsilon$  from quantities that are order-one with respect to  $\varepsilon$ .*

*To specify bounds with respect to  $\varepsilon$ , we sometimes make use of the shorthand notation  $\ll$ ,  $\lesssim$ , and  $\lesssim$ . To explain: We write*

$$A_\varepsilon \ll B_\varepsilon$$

*if for every  $C < \infty$ , we have  $A_\varepsilon/B_\varepsilon \leq 1/C$  for  $\varepsilon$  sufficiently small.*

*We write*

$$A_\varepsilon \lesssim B_\varepsilon$$

*if there exists a universal constant  $C < \infty$  such that  $A_\varepsilon \leq C B_\varepsilon$ , and similarly for  $A_\varepsilon \gtrsim B_\varepsilon$ . If both inequalities hold, then we write  $A_\varepsilon \sim B_\varepsilon$ .*

*We write*

$$A_\varepsilon \lesssim B_\varepsilon$$

*if for every  $\alpha > 0$  we have  $A_\varepsilon \leq B_\varepsilon + \alpha$  for  $\varepsilon$  sufficiently small, and similarly for  $A_\varepsilon \gtrsim B_\varepsilon$ . If both inequalities hold, then we write  $A_\varepsilon \approx B_\varepsilon$ .*

*We use numbered constants  $C_1, C_2$ , et cetera, to denote specific constants that we refer to later in the paper. On the other hand, we use  $C$  to denote a generic order-one constant whose value may change from place to place. Throughout the article,  $C$  or a numbered constant  $C_i$  is a constant that is universal except for a possible dependence on the potential  $V$ .*

We are now ready to state our results.

**1.2. Main results.** Recall that the boundary conditions imply that there must be at least one up layer and that any additional layers come in pairs. We will always consider the regime where the system size  $L_\varepsilon$  satisfies

$$1 \ll L_\varepsilon \lesssim \exp\left(\frac{c'_0}{\varepsilon}\right) \quad \text{for some } c'_0 < c_0. \quad (1.10)$$

(Recall that  $c_0$  is the energy cost defined in (1.8).) This is the regime in which one expects the probability of extra transitions to go to zero and in particular to obey the energetic and entropic scaling expressed in (1.2). Our first result captures this behavior on the exponential level.

**Theorem 1.5.** *Suppose that  $L_\varepsilon$  satisfies (1.10). Then for every  $n \in \mathbb{N}$  and  $\gamma > 0$  sufficiently small, there exists  $\varepsilon_0 > 0$  such that for  $\varepsilon \leq \varepsilon_0$ , one has the upper bound*

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \\ \leq (L_\varepsilon)^{2n} \exp\left(-\frac{2nc_0 - \gamma}{\varepsilon}\right), \end{aligned}$$

and the lower bound

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \\ \geq (L_\varepsilon)^{2n} \exp\left(-\frac{2nc_0 + \gamma}{\varepsilon}\right). \end{aligned}$$

*Remark 1.6.* One should note that because of the error term  $\gamma$ , our result sees only information on the exponential level. In particular, if one has an exponential system size such that

$$\varepsilon \log L_\varepsilon \approx c'_0 < c_0,$$

then what our result says is that for any  $n \in \mathbb{N}$  we have

$$\varepsilon \log \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \approx -2n(c_0 - c'_0).$$

*Remark 1.7.* Throughout the paper, when we say “ $u$  has  $2n+1$  layers,” we mean that  $u$  has at least  $2n+1$  layers.

*Remark 1.8.* As mentioned above, our techniques can also handle different boundary conditions, e.g., periodic boundary conditions or Dirichlet boundary conditions that do not enforce a transition layer. For instance, for periodic boundary conditions or Dirichlet conditions  $u(\pm L_\varepsilon) = 1$ , the probability of  $2n$  transition layers is bounded above and below by

$$(L_\varepsilon)^{2n} \exp\left(-\frac{2nc_0 \mp \gamma}{\varepsilon}\right),$$

respectively, while for homogeneous Dirichlet boundary conditions, the probability of  $n$  transition layers is bounded above and below by

$$(L_\varepsilon)^n \exp\left(-\frac{nc_0 \mp \gamma}{\varepsilon}\right),$$

respectively.

Our second main result states that, on scales larger than logarithmic in  $1/\varepsilon$ , the layer location is uniformly distributed in the following sense.

**Theorem 1.9.** Consider  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  in the regime

$$|\log \varepsilon| \ll L_\varepsilon \lesssim \exp\left(\frac{c'_0}{\varepsilon}\right) \quad \text{for some } c'_0 < c_0. \quad (1.11)$$

Let  $d_\varepsilon > 0$  be such that

$$|\log \varepsilon| \ll d_\varepsilon \leq L_\varepsilon.$$

Then uniformly for any  $x$  such that  $[x - d_\varepsilon, x + d_\varepsilon] \subseteq [-L_\varepsilon, L_\varepsilon]$ , we have

$$\frac{L_\varepsilon}{d_\varepsilon} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\text{there is an up layer contained in } [x - d_\varepsilon, x + d_\varepsilon]) \approx 1. \quad (1.12)$$

The theorem says that the probability of finding an up transition layer in a subinterval of length  $2d_\varepsilon$  given a system size  $2L_\varepsilon$  is approximately  $d_\varepsilon/L_\varepsilon$  in the sense expressed in (1.12), independent of the location of the subinterval. (The existence of an up transition layer somewhere in the system is forced by the boundary conditions.) In this sense, the layer locations are approximately uniformly distributed. The theorem is strongest when considering  $d_\varepsilon$  at the lower range of validity: It shows that the uniform distribution holds not only on macroscopic intervals but also down to the logarithmic scale.

We remark that the uniform distribution of the layer location in our regime is very different from the characterization of the layer distribution in the case  $L_\varepsilon = |\log \varepsilon|/4$  studied in [BBB08b]; see Subsection 1.4 below for more discussion.

**1.3. Methods: Markovianity, compact sets, and reflections.** Our approach for Theorem 1.5 relies on a simple idea. Namely, while we cannot use large deviation theory directly on  $(-L_\varepsilon, L_\varepsilon)$ , we can use the Markovianity of the underlying reference measure to reduce to order-one subintervals on which we can. In particular, by taking large (but order-one) subintervals and conditioning on the boundary values of a larger, surrounding subinterval, we can take advantage of *large deviation bounds* with a cost that is *to leading order independent of the subinterval size*. This method is similar in spirit to Freidlin and Wentzell's approach of calculating the expected exit time from a metastable domain for a diffusion process with small noise ([FW98], see Subsection 1.4 for a more detailed account of the related literature).

To illustrate the idea, suppose that we want to estimate the probability that there is a transition layer contained within  $[-\ell, \ell]$  for some  $\ell$  large. (Transition layers are introduced in Definition 1.3 above; roughly, they are layers connecting  $\pm 1$ .) The Markov property (Lemma 3.2) implies that this probability can be written as

$$\begin{aligned} & \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\text{transition in } (-\ell, \ell)) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu(du_-, du_+) \mu_{\varepsilon,(-2\ell,2\ell)}^{u_-,u_+}(\text{transition in } (-\ell, \ell)). \end{aligned} \quad (1.13)$$

Here  $\nu$  denotes the marginal distribution of the pair  $(u(-2\ell), u(2\ell))$ , and  $\mu_{\varepsilon,(-2\ell,2\ell)}^{u_-,u_+}$  denotes the distribution of paths on  $(-2\ell, 2\ell)$  with boundary conditions  $u_\pm$  (see Section 3 for a precise definition of this measure).

In Subsection 3.2 we establish large deviation estimates for the measures  $\mu_{\varepsilon,(-2\ell,2\ell)}^{u_-,u_+}$  that hold locally uniformly in the boundary values  $u_\pm$ . Hence for  $u_\pm$  in some large compact set, we can integrate over these bounds in (1.13). On the other hand, the probability that the boundary values  $u_\pm$  fall outside of the compact set  $[-M, M]$  for  $M \gg 1$  decays exponentially with  $M$  (see Lemma 4.1 below).



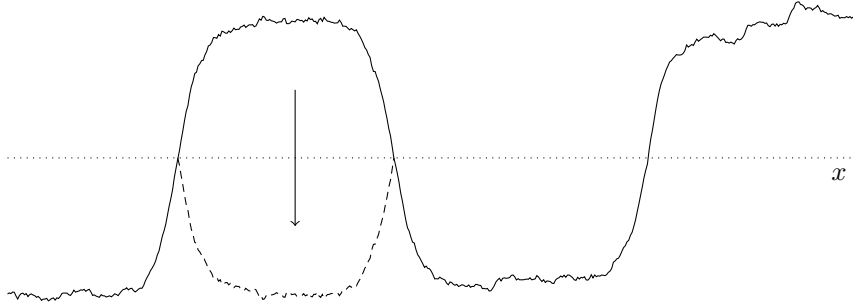


FIGURE 1.1. A vertical reflection turns a transition layer into a “wasted excursion” in which (roughly speaking) the path goes from  $-1$  to  $0$  and then back to  $-1$ . The probability of a wasted excursion on  $(-\ell, \ell)$  is approximately *independent* of the boundary conditions at  $\pm 2\ell$ .

For boundary values within the compact set  $[-M, M]$ , large deviation theory gives the uniform estimate

$$\begin{aligned} \mu_{\varepsilon, (-2\ell, 2\ell)}^{u_-, u_+}(\text{transition in } (-\ell, \ell)) \\ = \exp\left(-\frac{1}{\varepsilon}(\Delta E(\text{transition}) + o(1))\right). \end{aligned}$$

Here  $\Delta E(\text{transition})$  denotes the difference between the minimal energy of paths that perform a transition in  $(-\ell, \ell)$  and the minimal energy of any profile  $u$  that satisfies the boundary conditions  $u(\pm 2\ell) = u_{\pm}$ . (See Subsection 3.2 for a more complete discussion.)

Now we arrive at the second problem, which is more subtle. The issue is that the energy difference  $\Delta E(\text{transition})$  depends strongly on the boundary conditions. The cost that we are expecting to recover is  $c_0$ , defined in (1.8). However, if  $u_- \approx -1$  and  $u_+ \approx 1$ , for instance, then the energy *difference* is approximately zero! In this case, the information about the probability of a transition is encoded in the distribution  $\nu$ .

Our idea to handle the problem of dependence on the boundary conditions relies on Markovianity and the global symmetries of  $\mu_{\varepsilon, (-L\varepsilon, L\varepsilon)}^{-1, 1}$ . What we want to do is to transform a transition event into an event that does not feel the influence of the boundary conditions. Roughly, the new event will be that there are points  $x < y < z \in (-\ell, \ell)$  such that  $u(x) \approx u(z) \approx -1$  while  $u(y) = 0$ . (See Figure 1.1 for an illustration and Definitions 2.4 and 2.7 for formal definitions of these “wasted excursions.”) The expected cost for such an event is also  $c_0$ , and a little thought reveals that this should be the energy difference *regardless of the boundary conditions at  $\pm 2\ell$* . (For a result in this direction, see Lemma 2.5.)

In order to transform transitions into wasted excursions, we use the strong Markov property (see Lemma 3.3) and the symmetry of  $V$ . Specifically, we reflect paths vertically between certain hitting points of zero in such a way that leaves  $\mu_{\varepsilon, (-L\varepsilon, L\varepsilon)}^{-1, 1}$  invariant. For details, see for instance (4.22) and the subsequent calculations in the proof of Theorem 1.5.

A different reflection operator turns out to be useful when we come to the proof of the uniform distribution of the layer location in Theorem 1.9. Again

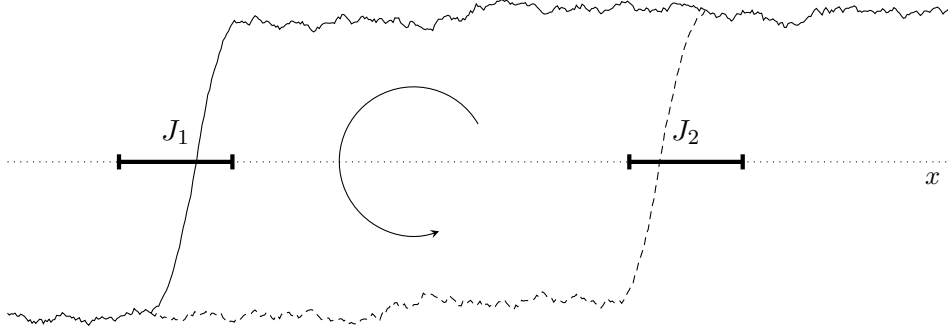


FIGURE 1.2. A point reflection between a hitting point of  $-1$  and a hitting point of  $+1$  moves the transition from the interval  $J_1$  into the interval  $J_2$ . As the point reflection preserves the measure, both events have the same probability.

the Markovianity and the symmetry of  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  are crucial. Here the rough idea is to show that the probability of finding the transition layer in any interval  $[y - d_\varepsilon, y + d_\varepsilon]$  is approximately the same as that of finding the layer in any other interval  $[z - d_\varepsilon, z + d_\varepsilon]$ . In Section 5, we construct a measure-preserving reflection operator that transforms paths with a transition in  $[y - d_\varepsilon, y + d_\varepsilon]$  into paths with a transition in (or near)  $[z - d_\varepsilon, z + d_\varepsilon]$ . We build this reflection operator using certain hitting points of  $-1$  and  $+1$  to the left and right of the transition layer. (This is illustrated in Figure 1.2.) Hence a key point is to prove that, on the set of paths with a transition in  $[y - d_\varepsilon, y + d_\varepsilon]$ , such hitting points exist with high probability. This fact is developed in Lemmas 5.1 and 5.2 using an iterated rescaling argument and large deviation bounds.

**1.4. Background literature and related results.** The study of the effect of a small noise on a physical system has a rich history in the chemistry, physics, and mathematics literature. With roots in the fluctuation theory of Einstein [Ei10] (1910), the path integral formulations of Wiener [W30] (1930) and Feynman [Fe48] (1948) lie at the heart of the large deviation theory for diffusion processes and the characterization of the corresponding invariant measure. One of the aspects to receive the most applied interest and significant mathematical attention is the question of the first exit time from a metastable basin. The exponential dependence of the mean exit time on the energy barrier goes back to Van't Hoff and Arrhenius [VH84, A89] (1889). Refining this picture, the so-called Kramers formula determines the prefactor in terms of the curvature of the potential at the critical points and was made famous in the 1940 paper by Kramers [Kr40], although the result (for the overdamped dynamics) had been derived as early as 1927 by Farkas [Fa27]. See the review paper by Hänggi, Talkner, and Borkovec [HTB90] for a thorough historical survey. The higher dimensional case was analyzed by Landauer & Swanson [LS61] in 1961 and further pursued by Langer (see for instance [L69], 1969).

In the mathematics literature, metastability for diffusion processes that depend only on time (i.e., constant in space) was explored early on in the paper by Pontryagin, Andronov, and Vitt [PAV33] (1933). The mathematical theory of large deviations was subsequently developed in the 1970s in papers by Wentzell and Freidlin (see for instance [WF70]) and Kifer [Ki74], and a landmark text is the

book of Freidlin-Wentzell [FW98] (published in Russian in 1979 and first published in English in 1984). On the level of the mean exit time, the Freidlin-Wentzell theory confirmed the exponential factor in the Kramers formula. The prefactor in Kramers' law for  $d > 1$  was established via formal asymptotic expansions in the famous paper by Matkowsky and Schuss [MS77] in 1977. A rigorous derivation was given by Sugiura in [S95] and independently and with a different method by Bovier, Eckhoff, Gaynard, and Klein [BEGK04, BGK05].

The small noise problem for stochastic *partial* differential equations appears more recently in the mathematics community. A seminal paper in extending the Freidlin-Wentzell theory to spatially varying diffusions is the paper of Faris and Jona-Lasinio [FJ82] from 1982, which specifically established and studied the action functional of the stochastic Allen-Cahn differential equation on a bounded system  $[0, L]$ . The invariant measure of stochastically perturbed reaction diffusion systems (including the Allen-Cahn equation) on a bounded domain is studied by Freidlin in [Fr88] in 1988. Recently, Barret, Bovier, and Méléard [BBM10, B12] and Berglund and Gentz [BG12] have established the mean exit time estimate including the prefactor for a class of equations including the Allen-Cahn equation.

As we have emphasized in the beginning of the introduction, in this paper we are concerned with the interplay between *small noise* and *large domain size*. Specifically, we are interested in system sizes that are exponential with respect to the inverse noise strength. Before turning to the invariant measure for unbounded systems, we remark that there is already an *entropic, system-size dependent component* of the mean switching time when there is a “flat” or “degenerate” saddle point, e.g., for the Allen-Cahn equation in the periodic case. Specifically, the prefactor picks up a factor that is proportional to the volume of the degenerate set. This fact was observed already by Glasstone, Laidler, and Eyring [GLE41] (1941) in the context of transition state theory, and the estimates in the setting of overdamped diffusions were developed by Langer [L69] (1969) and Matkowsky & Schuss [MS77] (1977). See also [VW08] for an independent, also formal, derivation.

The dynamics of the stochastic Allen-Cahn equation (1.1) have been considered by several authors. In particular, in the groundbreaking works of Funaki [Fu95] and Brascasse, De Masi, and Presutti [BDMP95], the dynamics of very similar equations were studied. In [Fu95], the equation (1.1) is considered on the whole line with boundary conditions that enforce one transition. The noise term  $\sqrt{2\varepsilon}\eta$  is multiplied by a function with compact support. In terms of our notation, the noise acts on an interval  $L_\varepsilon$  of length polynomial in  $\varepsilon^{-1}$ . In [BDMP95], the equation (1.1) is considered for  $L_\varepsilon = \varepsilon^{-1}$  with Neumann boundary conditions. In both articles, the initial condition is chosen close to the optimal profile of a single transition, and it is shown that the solution stays close to an optimal profile on timescales that are polynomial in  $\varepsilon^{-1}$ . The evolution of the midpoint of the transition layer is also characterized: In [Fu95], the interface dynamic is given by a stochastic differential equation that reflects the spatially dependent noise strength. In [BDMP95], it is shown that the midpoint performs a Brownian motion. The dynamic behavior observed in both of these articles is consistent with our results on the invariant measure. In particular, the Brownian motion of interfaces is consistent with the uniform distribution of layer location that we observe in Theorem 1.9.

Now let us consider the interplay between small noise and large domain size. The idea of understanding large deviation events on large spatial systems via a

decomposition into subintervals (intermediate in size between the logarithmic and exponential scale) is used in the paper [VW08] to heuristically derive the nucleation and propagation dynamics in the setting of an unequal-well potential. In rigorous work on the invariant measure for the equal-well case, the second author derived a concentration result for the measures  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  in [Web10] for system sizes that are large but algebraically bounded: specifically,  $L_\varepsilon \leq \varepsilon^{-\alpha}$  for  $\alpha < 2/3$ . The technique used there is completely different from the one employed in the present article, however. In [Web10], the measure is discretized to make rigorous the heuristic intuition that  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  is a Gibbs measure. Explicit bounds on the energy landscape and Gaussian concentration inequalities are then used to derive bounds on this discretized measure. This technique does not appear to be applicable for longer intervals because the discretization errors become too large.

In the articles [BBB08a] and [BBB08b], the special case of intervals growing like  $L_\varepsilon = \frac{1}{4}|\log \varepsilon|$  is studied. (The prefactor  $1/4$  depends on a specific choice of double-well potential.) The article uses the fact pointed out in [RVE05] that the measure  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  can be realized as the distribution of a diffusion process

$$du(x) = a_\varepsilon(u(x)) dx + \varepsilon^{1/2} dw(x) \quad u(-L_\varepsilon) = -1, \quad (1.14)$$

conditioned on the event  $u(L_\varepsilon) = 1$ . The drift term  $a_\varepsilon$  is the logarithmic derivative of the ground state of the Schrödinger operator  $-\varepsilon^2 \Delta + V$ . (In most cases, the drift  $a_\varepsilon$  cannot be given explicitly.) This is the extension to bounded intervals of the well-known equivalence for the measure on the real line, cf. [S79].

Building on the connection between the invariant measure of the PDE and the process in (1.14), [BBB08b] derives a concentration result around the one-parameter family of energy minimizers. Furthermore, the authors characterize the asymptotic distribution of the position of the interfacial layer. It is nonuniform due to the energetic repulsion from the boundary of the interval. To see this nonuniformity, the moderate scaling  $L_\varepsilon \approx |\log \varepsilon|$  is necessary. Incidentally, this shows that our lower bound  $d_\varepsilon \gg |\log \varepsilon|$  in Theorem 1.9 is optimal: Below the scale of  $|\log \varepsilon|$ , nonuniformity occurs. Loosely speaking, the results in [BBB08b] and ours are complimentary. They obtain finer results on logarithmically large system sizes, we obtain coarser results on exponentially large system sizes.

Results similar to (but different from) ours were obtained in [COP93] for a one-dimensional Ising model with ferromagnetic Kac potential. This is a spin model whose spins interact not only with their nearest neighbors, but with all spins in a given range. The authors study the limit in which this range diverges. This corresponds to the limit  $\varepsilon \downarrow 0$  that we investigate. Their main argument relies on a large deviation statement for the whole system in a *local topology*. This large deviation result implies, for example, that the local spin averages concentrate around  $\pm 1$  and that probability to see a transition from  $-1$  to  $+1$  in any given compact interval is exponentially small. The exponential rate is given by the energetic cost of a transition (similar to the constant  $c_0$  in this work). The significant difference between their large deviation bounds and ours is the dependence on the boundary condition. Their bounds state that the exponential decay of the probability of observing a certain behavior on an order-one interval is governed by the energy. We only get bounds for the measures conditioned on the boundary values on that interval. The difference is easy to appreciate on the level of the results. As mentioned,

the probability of seeing a transition on a given order-one subinterval in their setting is exponentially small, while—because of our boundary conditions—a similar statement cannot possibly hold in our case: Indeed, if it were to hold, we could sum over order-one subintervals and deduce that the probability to see a transition in the full system goes to zero with the noise, while in fact it is identically equal to one.

Finally, let us touch on the appearance of measures similar to  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  in the study of Schrödinger operators. The Feynman-Kac formula gives a way to solve the imaginary time Schrödinger equation (i.e., the heat equation with a potential) in terms of measures that are absolutely continuous with respect to Wiener measure. In this context, our model is often referred to as the  $\phi_1^4$  model and the limit  $\varepsilon \downarrow 0$  corresponds to the *semiclassical limit* in which the Planck constant  $\hbar$  is sent to zero. Lemma 4.1, for instance, is closely related (but not equivalent to) a statement about the decay of the ground state for the Schrödinger operator  $\varepsilon^2 \Delta + V$  as  $\varepsilon \downarrow 0$ .

**1.5. Organization.** We begin with preliminaries: In Section 2 we collect some properties of the energy functional, and in Section 3 we collect some probabilistic properties of  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  and of the underlying Gaussian measures. With these preliminaries in hand, we turn in Section 4 to the proof of our first result, Theorem 1.5. In Section 5 we prove Theorem 1.9, the uniform distribution of the layer location. Finally, in Section 6 we prove the various technical lemmas that have been used in support of the main theorems.

## 2. DETERMINISTIC PRELIMINARIES

In this section we discuss some more details about the energy functional  $E$  (cf. (1.6)). Our goal is to familiarize the reader with the common intuition about this energy, as well as to present some facts that will guide our method and appear later in proofs.

As described above, the potential term in the energy favors the states  $\pm 1$  and the gradient term in the energy leads to an energetic cost for transitions between these states. Given our large system and the boundary conditions (1.4), it is natural to consider the problem

$$\inf \{E_{(-\infty,\infty)}(u) : u(\pm\infty) = \pm 1\}.$$

As we mentioned, the minimum cost  $c_0$  can be calculated explicitly (cf. (1.9)). The calculations underlying this fact appear repeatedly in the proofs of our energy lemmas, so we begin by recalling them. The so-called Modica-Mortola trick (cf. [MM77]) uses the elementary inequality  $a^2 + b^2 \geq 2ab$  to observe:

$$\begin{aligned} & \inf \{E_{(-\infty,\infty)}(u) : u(\pm\infty) = \pm 1\} \\ &= \inf \left\{ \int_{-\infty}^{\infty} \left( \frac{1}{2} (\partial_x u)^2 + V(u) \right) dx : u(\pm\infty) = \pm 1 \right\} \\ &\geq \inf \left\{ \int_{-\infty}^{\infty} \sqrt{2V(u)} (\partial_x u) dx : u(\pm\infty) = \pm 1 \right\} \\ &= \int_{-1}^1 \sqrt{2V(u)} dx, \end{aligned}$$

which gives a lower bound on the energetic cost. For the matching upper bound, one observes that the equality  $a^2 + b^2 = 2ab$  holds if and only if  $a = b$ , so that the

minimum energetic cost is achieved precisely when

$$|\partial_x u| = \sqrt{2V(u)}. \quad (2.1)$$

For our boundary conditions, it is easy to see that the minimum is achieved for the strictly increasing function that satisfies

$$\partial_x u = \sqrt{2V(u)}. \quad (2.2)$$

We denote by  $m$  the minimizer that is normalized so that  $m(0) = 0$ . This function  $m$  is then the unique, centered, stationary solution of the Allen-Cahn equation on  $\mathbb{R}$  subject to the given boundary conditions, i.e., the solution of

$$\partial_x^2 m - V'(m) = 0 \quad m(0) = 0 \quad \text{and} \quad m(\pm\infty) = \pm 1.$$

In the case of the standard double-well potential  $V(u) = (1 - u^2)^2/4$ , one has  $m(x) = \tanh(x/\sqrt{2})$ .

For general potentials satisfying Assumption 1.1, the energy minimizer has similar qualitative properties to the hyperbolic tangent. In particular, what will be important for us is that the minimizer converges exponentially to  $\pm 1$  as  $x \rightarrow \pm\infty$ .

**Lemma 2.1** (Exponential decay of minimizer). *Under Assumption 1.1 on the potential  $V$ , there exists  $C < \infty$  such that the global energy minimizer  $m$  satisfies*

$$|m(x) - \text{sign}(x)| \leq C \exp\left(-\sqrt{\frac{V''(1)}{2}} |x|\right).$$

The exponential convergence to  $\pm 1$  follows directly from (2.2) and the quadratic behavior of  $V$  near the minima (cf., Assumption 1.1).

In addition to the exponential convergence to  $\pm 1$ , we see from (2.2) and Assumption 1.1 that outside of a neighborhood of  $\pm 1$ , the slope of  $m$  is bounded away from zero. Consequently, there is a characteristic length-scale associated to a transition layer. We will use this length-scale in an essential way. That is, since we cannot apply large deviation theory on the full system scale  $L_\varepsilon$ , we will decompose into subsystems of bounded size, typically called  $2\ell$  or  $4\ell$ . We will choose the subsystem size so that (with very large probability) a typical transition layer fits inside, which requires  $\ell$  to be large. In order to make these ideas precise, we begin by introducing the idea of a  $\delta^-$  transition layer. Simply put, instead of connecting  $\pm 1$ , it connects  $-1 + \delta$  with  $1 - \delta$ .

**Definition 2.2** ( $\delta^-$  transition layer). *Fix  $\delta \in (0, 1/2)$  and suppose  $x_- < x_+$ . We say that  $u$  has a  $\delta^-$  up transition layer between  $x_-$  and  $x_+$  if*

$$u(x_\pm) = \pm(1 - \delta) \quad \text{and} \quad |u(x)| < 1 - \delta \quad \text{for all } x \in (x_-, x_+).$$

*We say that  $u$  has a  $\delta^-$  down transition layer on  $(x_-, x_+)$  if the same condition holds true with signs reversed, and that  $u$  has a  $\delta^-$  transition layer if it has a  $\delta^-$  up or a  $\delta^-$  down transition layer.*

Since it is of course true that

$$\begin{aligned} & \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \\ & \leq \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \delta^- \text{ transition layers}), \end{aligned}$$

the proof of the upper bound in Theorem 1.5 will be established if we can show that for any  $\gamma > 0$  and for sufficiently small  $\delta > 0$ , there is an  $\varepsilon_0 > 0$  such that, for all  $\varepsilon \leq \varepsilon_0$ , we have

$$\begin{aligned} & \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \delta^- \text{ transition layers}) \\ & \lesssim (L_\varepsilon)^{2n} \exp\left(-\frac{2nc_0 - \gamma}{\varepsilon}\right). \end{aligned} \quad (2.3)$$

The main ingredient for establishing (2.3) is the uniform large deviation estimate from Proposition 3.4, below, which essentially reduces the problem to one of *energy estimates*. We will control the energy of suitable classes of functions up to a small  $\delta$ -dependence and ultimately absorb this error term into the large deviation error  $\gamma$  from the proposition.

One of the first steps will be to understand the length-scale associated to  $\delta^-$  transition layers. For any  $\delta \in (0, 1/2)$ , the optimal transition layer captured by the energy minimizer  $m$  goes from  $-1 + \delta$  to  $1 - \delta$  over a finite length-scale, and “typical layers” perform the transition on a similar length-scale. A question that we will have to address is how likely it is for a transition to take unusually long to complete a  $\delta^-$  transition. In the following lemma, we show that the difference of energies expressed in Proposition 3.4 is large for functions that perform unusually long transitions (uniformly with respect to the boundary values).

**Lemma 2.3** (Long transitions). *There exists a  $C_1 < \infty$  (depending only on  $V$ ) such that, for any  $M < \infty$  and any  $\delta \in (0, 1/2)$ , there exists an  $\ell_* < \infty$  with the following property. For any  $\ell \geq \ell_*$  and  $u_\pm \in [-M, M]$ , set*

$$\begin{aligned} \mathcal{A}^{\text{bc}} &:= \{u \in C([-2\ell, 2\ell]) : u(-2\ell) = u_- \text{ and } u(2\ell) = u_+\}, \\ \mathcal{A}_0^{\text{bc}} &:= \{u \in \mathcal{A}^{\text{bc}} : \text{for all } x \in [-\ell, \ell], u(x) \in [-1 + \delta, 1 - \delta]\}. \end{aligned}$$

Then we have

$$\inf_{u \in \mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \geq \frac{2\delta^2 \ell}{C_1}. \quad (2.4)$$

The proof of Lemma 2.3 is given in Subsection 6.1. This lemma together with the large deviation bound from Proposition 3.4 will imply that for  $\gamma$  small with respect to  $\delta^2 \ell$ , the probability of finding such a layer is bounded above by

$$\exp\left(-\frac{2\delta^2 \ell / C_1 - \gamma}{\varepsilon}\right) \leq \exp\left(-\frac{\delta^2 \ell}{C_1 \varepsilon}\right),$$

which we can make negligible by choosing  $\ell$  sufficiently large.

Now we would like to show that the exponential factor in the probability of finding a  $\delta^-$  layer is close to  $c_0$ , defined in (1.8). Specifically, we expect it to be approximately

$$\int_{-1+\delta}^{1-\delta} \sqrt{2V(s)} ds.$$

The problem, which we already alluded to at the end of Subsection 1.3, is that the boundary values (for instance  $u(-2\ell) \approx -1$ ,  $u(2\ell) \approx 1$ ) may make it likely to find a layer. Hence, we will employ reflection operators to transform  $\delta^-$  transition layers into events that are unlikely *regardless of the boundary conditions*. We will call such events wasted  $\delta^-$  excursions:

**Definition 2.4** (Wasted  $\delta^-$  excursion). *For any  $\delta \in (0, 1/2)$ , we will say that  $u$  has a wasted  $\delta^-$  excursion on  $(-\ell, \ell)$  if there exist points*

$$-\ell \leq x_- < x_0 < x_+ \leq \ell$$

*such that*

$$|u(x_0)| \leq \delta$$

*and*

$$\text{either } |u(x_\pm) - 1| \leq \delta \quad \text{or} \quad |u(x_\pm) + 1| \leq \delta.$$

As described above for long transitions, we will estimate the probability of such events using the large deviation estimate from Proposition 3.4. We note that the proposition requires minimizing energy over a ball (in the space of continuous functions) around the set of interest. Because of the way we have defined wasted excursions, a ball of radius  $\delta$  around the set of functions with a  $\delta^-$  excursion in a given interval is equal to the set of functions with a  $(2\delta)^-$  excursion in that interval. Hence, our large deviation estimate together with an energetic estimate will bound the probability that we are after. The following lemma contains the necessary energetic estimate: namely, that the difference of energies described in our large deviation estimate is bounded below by  $c_0$  plus a small term.

**Lemma 2.5.** *There exists a constant  $C < \infty$  such that for every  $M < \infty$  and  $\delta \in (0, 1/2)$ , there exists a constant  $\ell_* < \infty$  with the following property. For any  $\ell \geq \ell_*$  and any boundary conditions  $u_\pm \in [-M, M]$ , set*

$$\begin{aligned} \mathcal{A}^{\text{bc}} &:= \{u \in C([-2\ell, 2\ell]): u(\pm 2\ell) = u_\pm\}, \\ \mathcal{A}_0^{\text{bc}} &:= \{u \in \mathcal{A}^{\text{bc}}: u \text{ has a wasted } \delta^- \text{ excursion in } (-\ell, \ell)\}. \end{aligned}$$

*Define the optimal cost*

$$c_\ell := \inf_{\mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u). \quad (2.5)$$

*Then we have*

$$c_\ell - c_0 \geq -C\delta. \quad (2.6)$$

The proof of Lemma 2.5 is given in Subsection 6.1. It gives us the exponential factor in the desired estimate (2.3), above.

For the lower bound in Theorem 1.5, we will work with so-called  $\delta^+$  transition layers between  $-1 - \delta$  and  $1 + \delta$ .

**Definition 2.6** ( $\delta^+$  transition layer). *Fix  $\delta \in (0, 1/2)$ . We say that  $u$  has a  $\delta^+$  up transition layer within the interval  $(-\ell, \ell)$  if there exist points*

$$-\ell \leq x_- < x_+ \leq \ell$$

*such that*

$$u(x_\pm) = \pm(1 + \delta).$$

*We say that  $u$  has a  $\delta^+$  down transition layer on  $(-\ell, \ell)$  if the same condition holds true with signs reversed, and that  $u$  has a  $\delta^+$  transition layer if it has a  $\delta^+$  up or a  $\delta^+$  down transition layer.*



In analogy with the  $\delta^-$  transition layers that we use for the upper bound,  $\delta^+$  transition layers will be convenient for the lower bound. Since the probability of having  $(2n+1)$  transition layers is greater than the probability of having  $(2n+1)$   $\delta^+$  transition layers, it will suffice to show that

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \delta^+ \text{ transition layers}) \\ \gtrsim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 - \gamma}{\varepsilon} \right). \end{aligned}$$

We will establish this bound by reflecting in order to transform the  $\delta^+$  transition layers into some kind of “wasted excursions” whose probability we can bound, independently of the boundary conditions.

**Definition 2.7** (Wasted  $\delta^+$  excursion). *For any  $\delta \in (0, 1/2)$ , we will say that  $u$  has a wasted  $\delta^+$  excursion on  $(-\ell, \ell)$  if there exist points*

$$-\ell \leq x_- < x_0 < x_+ \leq \ell$$

*such that*

$$u(x_\pm) \leq -1 - \delta, \quad u(x_0) = 0.$$

(We will use only the wasted  $\delta^+$  excursions that come from below, but of course it would be straightforward to define the analogue with  $u(x_\pm) \geq 1 + \delta$ , and they would obey the same energetic and probabilistic bounds.)

As in the case of the upper bound, we need an energetic lemma that will control the contribution to the large deviation estimate for wasted  $\delta^+$  excursions. Because of the form of the large deviation estimate that we will develop in Section 3 (see Proposition 3.5 below), it will be convenient for us to introduce the energy bound on the following set of functions:

$$\begin{aligned} \mathcal{A}_{\delta, pre}^{bc} := \left\{ u \in \mathcal{A}^{bc} : \text{there exist points } -\ell \leq x_- < x_0 < x_+ \leq \ell \right. \\ \left. \text{with } u(x_-) \leq -1 - 2\delta, u(x_+) \leq -1 - 2\delta, u(x_0) \geq \delta \right\}. \end{aligned} \quad (2.7)$$

It is easy to see that a  $\delta$  ball (with respect to the sup norm) around  $\mathcal{A}_{\delta, pre}^{bc}$  is equal to the set of functions with wasted  $\delta^+$  excursions on  $(-\ell, \ell)$ . This fact is what will later be useful for the lower bound. For now, we record the following energetic fact, which plays the role for the lower bound that Lemma 2.5 played for the upper bound.

**Lemma 2.8.** *There exists a constant  $C < \infty$  such that for every  $M < \infty$  and  $\delta \in (0, 1/2)$ , there exists a constant  $\ell_* < \infty$  with the following property. For any  $\ell \geq \ell_*$  and  $u_\pm \in [-M, 0]$ , set*

$$\mathcal{A}^{bc} := \{u \in C([-2\ell, 2\ell]) : u(\pm 2\ell) = u_\pm\}$$

*and  $\mathcal{A}_{\delta, pre}^{bc}$  as above in (2.7).*

*Define the optimal cost*

$$c_\ell := \inf_{\mathcal{A}_{\delta, pre}^{bc}} E_{(-2\ell, 2\ell)}(u) - \inf_{\mathcal{A}^{bc}} E_{(-2\ell, 2\ell)}(u).$$

*Then we have*

$$c_\ell - c_0 \leq C \delta.$$

We will need to consider some additional properties of the energy as we prove the main theorems, but we defer their discussion to a later time when their motivation and hypotheses will be clearer. With the central facts about the energy in hand, we now turn to the probabilistic background for our paper.

### 3. PROBABILISTIC PRELIMINARIES

In this section, we collect some probabilistic facts about the Gaussian measures  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  and the measures  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ . After stating a precise definition and some elementary symmetry properties, we will discuss Markov properties satisfied by these measures in Subsection 3.1 and large deviation bounds in Subsection 3.2.

For every  $x_- < x_+$ , we denote by  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  the distribution of a Brownian bridge with homogeneous boundary conditions on  $[x_-, x_+]$  whose variance is proportional to  $\varepsilon$ . To be more precise,  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  is the unique centered Gaussian measure on the space of continuous functions  $C([x_-, x_+])$  such that, for all  $x_1, x_2 \in [x_-, x_+]$ , one has

$$\begin{aligned} \mathbb{E}_{\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}} \left( u(x_1) u(x_2) \right) \\ = \frac{\varepsilon}{x_+ - x_-} \left( (x_1 - x_-)(x_+ - x_2) \wedge (x_2 - x_-)(x_+ - x_1) \right). \end{aligned} \quad (3.1)$$

Equivalently, one can say that  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  is the centered Gaussian measure whose Cameron-Martin space is given by the Sobolev space  $H_0^1([x_-, x_+])$  with vanishing boundary conditions equipped with the homogeneous scalar product

$$\frac{1}{\varepsilon} \int_{x_-}^{x_+} \partial_x u \partial_x v \, dx.$$

Indeed, the right-hand side of (3.1) is the Green's function for  $\frac{1}{\varepsilon} \partial_x^2$  with Dirichlet boundary conditions.

In the sequel, we often use the notation

$$I_{x_-, x_+}(u) := \frac{1}{2} \int_{x_-}^{x_+} (\partial_x u)^2 \, dx \quad (3.2)$$

to denote the Gaussian part of the energy of a function  $u$  on the interval  $(x_-, x_+)$ .

It is common to think of  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  as a Gibbs measure

$$\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0} \propto \exp \left( -\frac{1}{\varepsilon} I_{x_-, x_+}(u) \right) du \quad (3.3)$$

with energy  $I_{x_-, x_+}$  and noise strength  $\propto \varepsilon$ . Of course, (3.3) does not make rigorous sense because there is no “flat measure”  $du$  on path space, and  $I_{x_-, x_+}(u)$  is almost surely infinite under  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$ . The heuristic formula (3.3) is motivated by finite dimensional approximations and it gives the right intuition for the large deviation bounds.

For more general boundary conditions  $u_-, u_+ \in \mathbb{R}$ , we can define  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  as the image measure of  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  under the shift map

$$u(x) \mapsto u(x) + h_{(x_-, x_+)}^{u_-, u_+}(x),$$

where  $h$  is the affine function interpolating the boundary conditions:

$$h_{(x_-, x_+)}^{u_-, u_+}(x) := \frac{x - x_-}{x_+ - x_-} u_+ + \frac{x_+ - x}{x_+ - x_-} u_-. \quad (3.4)$$

Similarly to (1.5), for any choice of boundary condition  $u_{\pm}$  and on any interval  $(x_-, x_+)$ , we denote by  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  the probability measure whose density with respect to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  can be expressed as

$$\frac{d\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}}{d\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}}(u) = \frac{1}{\mathcal{Z}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}} \exp\left(-\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx\right). \quad (3.5)$$

Here we have introduced the notation

$$\mathcal{Z}_{\varepsilon, (x_-, x_+)}^{u_-, u_+} := \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}} \left( \exp\left(-\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx\right) \right)$$

for the normalization constant that ensures that  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  is indeed a probability measure.

As we have indicated in the introduction, there are symmetry properties of the measures  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  and  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  that will play an important role in our argument. Observe for example that both  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  and  $\mu_{\varepsilon, (x_-, x_+)}^{0,0}$  are invariant under the *vertical reflection*  $u \mapsto Ru$  and the *horizontal reflection*  $u \mapsto Su$  where

$$Ru(x) := -u(x) \quad \text{and} \quad Su(x) := u(x_+ + x_- - x).$$

Furthermore, the measures  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{-1,1}$  and  $\mu_{\varepsilon, (x_-, x_+)}^{-1,1}$  are invariant under the *point reflection*  $u \mapsto RSu$ .

**3.1. Markov properties.** We first present a two-sided version of the Markov property for the measures  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  and  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ , which states that for any fixed points  $x_- \leq \hat{x}_- < \hat{x}_+ \leq x_+$  and for  $u$  distributed according to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  (or  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ ), the conditional distribution of  $(u(x), x \in [\hat{x}_-, \hat{x}_+])$ , given all the information about  $u(x)$  for  $x \in [x_-, x_+] \setminus (\hat{x}_-, \hat{x}_+)$ , is  $\mathcal{W}_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{u(\hat{x}_-), u(\hat{x}_+)}$  (or  $\mu_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{u(\hat{x}_-), u(\hat{x}_+)}$ ). Then in Lemma 3.3, we give the *strong* Markov property, which states that the same statement holds true when the deterministic points  $\hat{x}_{\pm}$  are replaced by left and right stopping points  $\chi_{\pm}$ . The proofs of these statements are quite standard. For completeness, we have included them in Subsection 6.2.

In the case of the measures  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ , the Markov property can be stated in the following way. For  $\hat{x}_- < \hat{x}_+$ , we define the piecewise linearization  $u_{\hat{x}_-}^{\hat{x}_+}$  of  $u$  between  $\hat{x}_-$  and  $\hat{x}_+$  as

$$u_{\hat{x}_-}^{\hat{x}_+}(x) = \begin{cases} h_{(\hat{x}_-, \hat{x}_+)}^{u(\hat{x}_-), u(\hat{x}_+)}(x) & \text{if } x \in (\hat{x}_-, \hat{x}_+) \\ u(x) & \text{else.} \end{cases} \quad (3.6)$$

Recall the definition (3.4) of  $h_{(x_-, x_+)}^{u_-, u_+}$ . Then the following holds.

**Lemma 3.1.** *Suppose  $x_- \leq \hat{x}_- < \hat{x}_+ \leq x_+$  are fixed, non-random points. Then under  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  the random functions  $u - u_{\hat{x}_-}^{\hat{x}_+}$  and  $u_{\hat{x}_-}^{\hat{x}_+}$  are independent. Furthermore,  $u - u_{\hat{x}_-}^{\hat{x}_+}$  is zero outside of  $(\hat{x}_-, \hat{x}_+)$  and is distributed according to  $\mathcal{W}_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{0,0}$  between the two points.*

Due to the lack of spatial homogeneity, the corresponding property for the measures  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  has to be stated in a different way. For  $I \subseteq [x_-, x_+]$ , we denote by  $\mathcal{F}_I$  the sigma-algebra generated by  $u(x)$  for  $x \in I$ , completed with respect to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ .

We also introduce the following notation that extends the measures to paths on a larger domain by prescribing the values outside of an interval. Suppose that  $[\hat{x}_-, \hat{x}_+] \subseteq [x_-, x_+]$  and that  $\mathbf{u} \in C([x_-, x_+])$  is a fixed path. We say that  $u$  is distributed according to  $\mathcal{W}_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{\mathbf{u}}$ , resp.  $\mu_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{\mathbf{u}}$ , if it almost surely coincides with  $\mathbf{u}$  outside of the interval  $[\hat{x}_-, \hat{x}_+]$  and is distributed according to  $\mathcal{W}_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{\mathbf{u}(\hat{x}_-), \mathbf{u}(\hat{x}_+)}$ , resp.  $\mu_{\varepsilon, (\hat{x}_-, \hat{x}_+)}^{\mathbf{u}(\hat{x}_-), \mathbf{u}(\hat{x}_+)}$ , on  $[\hat{x}_-, \hat{x}_+]$ .

Then the Markov property takes the following form.

**Lemma 3.2.** *Suppose  $x_- \leq \hat{x}_- < \hat{x}_+ \leq x_+$  are fixed, non-random points. Then for any bounded measurable test function  $\Phi: C([x_-, x_+]) \rightarrow \mathbb{R}$ , we get the following identity:*

$$\mathbb{E}_{(x_-, x_+)}^{\mu_{\varepsilon, u_-, u_+}}(\Phi | \mathcal{F}_{[x_-, \hat{x}_-]} \vee \mathcal{F}_{[\hat{x}_+, x_+]}) = \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mu_{\varepsilon, \mathbf{u}}}(\Phi). \quad (3.7)$$

Here  $\mathcal{F}_{[x_-, \hat{x}_-]} \vee \mathcal{F}_{[\hat{x}_+, x_+]}$  denotes the smallest sigma-algebra that contains all sets in  $\mathcal{F}_{[x_-, \hat{x}_-]}$  and  $\mathcal{F}_{[\hat{x}_+, x_+]}$ .

We will typically use (3.7) in the following way: For given points  $x_- \leq x_1 \leq x_2 \leq \dots \leq x_{2n} \leq x_+$  and given events  $\mathcal{A}_i \in \mathcal{F}_{[x_{2i-1}, x_{2i}]}$ , we can write

$$\begin{aligned} & \mathbb{E}_{(x_-, x_+)}^{\mu_{\varepsilon, u_-, u_+}}(\mathbf{1}_{\mathcal{A}_1} \dots \mathbf{1}_{\mathcal{A}_n}) \\ &= \int_{\mathbb{R}^{2n}} \nu_{x_1, \dots, x_{2n}}(du_1, \dots, du_{2n}) \mathbb{E}_{(x_1, x_2)}^{\mu_{\varepsilon, u_1, u_2}}(\mathbf{1}_{\mathcal{A}_1}) \dots \mathbb{E}_{(x_{2n-1}, x_{2n})}^{\mu_{\varepsilon, u_{2n-1}, u_{2n}}}(\mathbf{1}_{\mathcal{A}_n}). \end{aligned} \quad (3.8)$$

Here  $\nu_{x_1, \dots, x_{2n}}$  denotes the distribution of the random vector  $(u(x_1), \dots, u(x_{2n}))$  under  $\mu_{\varepsilon, (x_-, x_+)}^{-1, 1}$ . Formula (3.8) follows directly by applying (3.7)  $n$  times.

To state the strong Markov property, we additionally need the notion of left and right stopping points. These are defined analogously to stopping times for Markov processes. A random variable  $\chi_-$  taking values in  $[x_-, x_+]$  will be called a *left stopping point* if for all  $x \in [x_-, x_+]$  the event  $\{\chi_- \leq x\}$  is contained in  $\mathcal{F}_{[x_-, x]}$ . In the same way a random variable  $\chi_+$  is called a *right stopping point* if for all  $x$  the event  $\{\chi_+ \geq x\}$  is contained in  $\mathcal{F}_{[x, x_+]}$ . In all of our applications the stopping points  $\chi_{\pm}$  are going to be left or rightmost hitting points of a closed set. It is easy to check that these random points are indeed left and right stopping points as defined above.

For given left and right stopping points  $\chi_{\pm}$ , we define the sigma-algebra  $\mathcal{F}_{[x_+, \chi_-]}$  of events that occur left of  $\chi_-$  and the sigma-algebra  $\mathcal{F}_{[\chi_+, x_+]}$  of events that happen to the right of  $\chi_+$  by

$$\begin{aligned} \mathcal{F}_{[x_-, \chi_-]} &:= \{\mathcal{A} \in \mathcal{F}_{[x_-, x_+]}: \forall x \quad \mathcal{A} \cap \{\chi_- \leq x\} \in \mathcal{F}_{[x_-, x]}\}, \\ \mathcal{F}_{[\chi_+, x_+]} &:= \{\mathcal{A} \in \mathcal{F}_{[x_-, x_+]}: \forall x \quad \mathcal{A} \cap \{\chi_+ \geq x\} \in \mathcal{F}_{[x, x_+]}\}. \end{aligned}$$

The strong Markov property can be stated in an analogous way to (3.7).

**Lemma 3.3.** *Suppose  $\chi_-$  and  $\chi_+$  are left and right stopping points with  $\chi_- < \chi_+$  almost surely. Suppose that  $\Phi: C([x_-, x_+]) \rightarrow \mathbb{R}$  is measurable and bounded.*

Then for any  $u_{\pm} \in \mathbb{R}$ , we get the following identities

$$\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon}, u_-, u_+} (\Phi | \mathcal{F}_{[x_-, \chi_-]} \vee \mathcal{F}_{[\chi_+, x_+]}) = \mathbb{E}_{(\chi_-, \chi_+)}^{\mathcal{W}_{\varepsilon}, \mathbf{u}} (\Phi) \quad (3.9)$$

and

$$\mathbb{E}_{(x_-, x_+)}^{\mu_{\varepsilon}, u_-, u_+} (\Phi | \mathcal{F}_{[x_-, \chi_-]} \vee \mathcal{F}_{[\chi_+, x_+]}) = \mathbb{E}_{(\chi_-, \chi_+)}^{\mu_{\varepsilon}, \mathbf{u}} (\Phi). \quad (3.10)$$

The strong Markov property is a crucial ingredient in the proofs of both Theorem 1.5 and Theorem 1.9. Let us illustrate how it is used in the proof of Theorem 1.5. Let  $\chi_-$  be the leftmost hitting of zero to the right of a given point  $x_-$  and  $\chi_+$  the rightmost hitting of zero to the left of a given point  $x_+$ . The values  $u(\chi_{\pm})$  in the formulas (3.9) and (3.10) are almost surely 0. Then we can use the invariance of  $\mathcal{W}_{\varepsilon, (\chi_-, \chi_+)}^{0,0}$  and  $\mu_{\varepsilon, (\chi_-, \chi_+)}^{0,0}$  under vertical reflection  $R$  to conclude that the whole right-hand side of (3.9) and (3.10) is invariant under vertical reflection on  $[\chi_-, \chi_+]$ . In Section 4, we will use this observation to reduce the problem of calculating the probability of transition layers to computing the probability of wasted excursions (see Definition 2.4).

**3.2. Large deviations.** Large deviation estimates for the measures  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  constitute an important ingredient for our argument. Large deviation bounds for Gaussian measures with a small variance, e.g., for  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ , are well-known (see e.g. [Bog98, Sec. 4.9]). They can be extended to the measures  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  with an “exponential tilting” argument (see e.g. [Var84], or [dH00, p.34]) in a standard way. Let  $\mathcal{A}^{\text{bc}}$  represent the set of continuous paths  $u$  on  $[x_-, x_+]$  that satisfy  $u(x_{\pm}) = u_{\pm}$ . The estimates then state that for every *closed* set  $\mathcal{A} \subseteq \mathcal{A}^{\text{bc}}$  and every  $\gamma > 0$ , there exists  $\varepsilon_0 > 0$  such that, for  $\varepsilon \leq \varepsilon_0$ , we have

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(\mathcal{A}) \leq \exp \left( -\frac{1}{\varepsilon} (\Delta E(\mathcal{A}) - \gamma) \right). \quad (3.11)$$

Similarly, for every *open* set  $\mathcal{A} \subseteq \mathcal{A}^{\text{bc}}$  and  $\gamma > 0$  there exists  $\varepsilon_0 > 0$  such that, for  $\varepsilon \leq \varepsilon_0$ , we have

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(\mathcal{A}) \geq \exp \left( -\frac{1}{\varepsilon} (\Delta E(\mathcal{A}) + \gamma) \right). \quad (3.12)$$

Here the energy difference  $\Delta E(\mathcal{A})$  is defined as

$$\Delta E(\mathcal{A}) := \inf_{u \in \mathcal{A}} E(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E(u). \quad (3.13)$$

Here and in the sequel, all topological notions like open and closed refer to the uniform topology, i.e., the topology generated by

$$\|u\|_{\infty} := \sup_{x \in [x_-, x_+]} |u(x)|. \quad (3.14)$$

Although we will not make use of it here, we remark that the bounds (3.11) and (3.12) are also true for different choices of topology. The Gaussian large deviation bounds hold for any separable Banach space that supports the Gaussian measure, and the “exponential tilting” works as soon as the exponential density is continuous.

A priori, the choice of  $\varepsilon_0$  depends not only on  $\gamma$  but also on the interval length  $\ell := x_+ - x_-$ , the boundary data  $u_{\pm}$ , and even the set  $\mathcal{A}$  itself. As pointed out in Subsection 1.3, however, our argument requires integrating probabilities for different boundary conditions. Therefore, we need to know that we can choose the same  $\varepsilon_0$  for these different boundary conditions simultaneously. Moreover, in

Lemma 5.1 we will need uniform estimates for measures with different potentials. Hence, we require uniform large deviation estimates, which is the content of the following two propositions. They deliver local uniformity with respect to  $\ell$ ,  $u_{\pm}$ ,  $\mathcal{A}$ , and even with respect to the potential function  $V$ . To state the result, it is convenient to introduce the notation

$$I_{x_{\pm}}^{u_{\pm}} := I(h_{(x_{-}, x_{+})}^{u_{-}, u_{+}}) = \frac{1}{2} \frac{(u_{+} - u_{-})^2}{x_{+} - x_{-}} \quad (3.15)$$

for the minimal Gaussian energy with the given boundary conditions. We will also write

$$B(\mathcal{A}, \delta) = \{u : \exists v \in \mathcal{A}, \|v - u\|_{\infty} \leq \delta\}$$

for the  $\delta$  neighborhood of a set  $\mathcal{A}$ .

**Proposition 3.4** (Large deviation upper bound). *Fix constants  $1 < M, R < \infty$  and  $0 < \ell_{-} < \ell_{+} < \infty$ . For any  $x_{\pm} \in \mathbb{R}$  with  $x_{+} - x_{-} \in [\ell_{-}, \ell_{+}]$  and any  $u_{\pm} \in [-M, M]$ , let  $\mathcal{A}$  be a measurable subset of  $C([x_{-}, x_{+}])$  consisting of paths  $u$  that satisfy the boundary conditions  $u(x_{\pm}) = u_{\pm}$ . Additionally, assume that*

$$\inf_{u \in \mathcal{A}} E(u) - I_{x_{\pm}}^{u_{\pm}} \leq R. \quad (3.16)$$

*Then for any  $\delta, \gamma > 0$  there exists an  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  we have*

$$\mu_{\varepsilon, (x_{-}, x_{+})}^{u_{-}, u_{+}}(\mathcal{A}) \leq \exp\left(-\frac{1}{\varepsilon}(\Delta E(B(\mathcal{A}, \delta)) - \gamma)\right), \quad (3.17)$$

*where  $\Delta E$  is defined in (3.13). This  $\varepsilon_0$  depends on  $M, R, \ell_{\pm}, \delta$ , and  $\gamma$  but not on the particular choice of  $x_{\pm}, u_{\pm}$ . It only depends on the set  $\mathcal{A}$  through the choice of  $R$  in condition (3.16). Furthermore,  $\varepsilon_0$  depends on  $V$  only through the local Lipschitz norm*

$$\sup_{|v| \leq M + \sqrt{2^{-1}(\ell_{+} R + 1)} + 1} |V'(v)|.$$

In particular, the same bounds hold for the same  $\varepsilon_0$  if  $V$  varies over a set of potentials with uniformly bounded local  $C^1$ -norm. This uniformity of (3.17) with respect to  $V$  will be used in Subsection 6.6. There, it will be applied to the family  $\{4^k V(2^{-k}(u - 1) + 1) : k \in \mathbb{N}\}$  of rescaled versions of  $V$ .

We also get the corresponding lower bounds without a condition on the minimal energy of  $E(u)$  for  $u \in \mathcal{A}$ .

**Proposition 3.5** (Large deviation lower bound). *Fix constants  $M$  and  $0 < \ell_{-} < \ell_{+} < \infty$ . Suppose that  $\ell = x_{+} - x_{-} \in [\ell_{-}, \ell_{+}]$  and  $u_{\pm} \in [-M, M]$ . Assume that there exists an energy minimizer*

$$u_{*} = \operatorname{argmin}_{u \in \mathcal{A}} E(u)$$

*satisfying  $u_{*} \in [-M, M]$ . Then, for any  $\gamma > 0$  and  $\delta > 0$  small enough, there exists  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  there holds*

$$\mu_{\varepsilon, (x_{-}, x_{+})}^{u_{-}, u_{+}}(B(\mathcal{A}, \delta)) \geq \exp\left(-\frac{1}{\varepsilon}(\Delta E(\mathcal{A}) + \gamma)\right), \quad (3.18)$$

*where  $\Delta E$  is defined in (3.13). As above,  $\varepsilon_0$  depends on  $M, \ell_{\pm}, \delta$ , and  $\gamma$ , but not on the particular choice of  $x_{\pm}, u_{\pm}$  or the set  $\mathcal{A}$ .*

The same remark about the uniform dependence on  $V$  holds for the lower bounds.

*Remark 3.6.* The existence of energy minimizers  $u_*$  in  $\mathcal{A}$  satisfying  $u_* \in [-M, M]$  is not necessary and it can be replaced by an approximation. Actually, we will show the Proposition under the slightly weaker assumption that for every  $\gamma > 0$  there exists a profile  $u_\gamma \in \mathcal{A}$  with  $uu_\gamma(x) \in [-M, M]$  for all  $x \in [x_-, x_+]$  and such that

$$E(u_\gamma) \leq \inf_{u \in \mathcal{A}} E(u) + \gamma. \quad (3.19)$$

The proofs of these Propositions are essentially a careful copy of the classical proofs and can be found in Subsection 6.3. Let us remark here that we do not expect the bounds (3.11) and (3.12) to hold uniformly for all open or closed sets. In fact, the argument for the classical statements makes use of qualitative properties such as existence of coverings by finitely many open sets. One sums over this finite number and uses the fact that, for  $\varepsilon$  small enough, only the largest summand matters. For different open or closed sets, this finite number will in general be different, and the choice of  $\varepsilon_0$  would also be different. We can resolve this issue by taking the  $\delta$  neighborhood of  $\mathcal{A}$  in the bounds (3.17) and (3.18) as a uniform version of the topological assumptions on  $\mathcal{A}$ .

#### 4. PROOFS OF THEOREM 1.5: DOMINATION BY SINGLE TRANSITION LAYER OF MINIMAL ENERGY

In this section we prove Theorem 1.5. This theorem estimates the exponentially small probability of having more than one layer (with the correct entropic effect and exponential factor). Hence, the most likely functions are those with only one transition layer.

As outlined in Subsection 1.3, at the heart of the method is the idea of decomposing the invariant measure into conditional measures and the corresponding marginals, so that we can reduce to estimating the probability of transition layers on order-one subintervals. When the boundary data of the subinterval falls within a compact set  $[-M, M]$ , large deviation theory will allow us to estimate probabilities in the spirit of

$$\int_{-M}^M \int_{-M}^M \nu(du_-, du_+) \mu_{\varepsilon, (-2\ell, 2\ell)}^{u_-, u_+} (\text{there is a transition layer in } (-\ell, \ell)).$$

On the other hand, the probability that  $|u(\pm 2\ell)| \geq M$  is uniformly small. Before turning to the proofs of the main theorems, we introduce this fact about the decay of the one-point distribution.

**Lemma 4.1.** *There exist  $M_1 < \infty$ ,  $C_2 < \infty$  (depending only on  $V$ ) such that the following holds. For any  $M \geq M_1$ , there exists  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  and any  $x_0$  in  $(-L_\varepsilon, L_\varepsilon)$  there holds*

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(|u(x_0)| \geq M) \leq \exp\left(-\frac{M}{\varepsilon C_2}\right). \quad (4.1)$$

The proof of the lemma is given in Subsection 6.4. With this preliminary estimate in hand, we turn now to the proof of Theorem 1.5. We consider separately the upper and lower bounds.

*Proof of Theorem 1.5.* Fix  $\gamma > 0$ . Fix a corresponding  $\delta > 0$  sufficiently small. Let  $\ell$  and  $M$  be large constants to be specified later. To begin with, let  $\ell$  be large enough so that (2.4) and (2.6) hold for the given  $\delta$ . We will divide the system into  $2N_\varepsilon$  intervals with

$$N_\varepsilon := \left\lfloor \frac{L_\varepsilon}{\ell} \right\rfloor, \quad (4.2)$$

labelling the endpoints:

$$x_{\pm k} := \begin{cases} \pm k \ell, & k = 0, \dots, (N_\varepsilon - 1), \\ \pm L_\varepsilon, & k = \pm N_\varepsilon. \end{cases} \quad (4.3)$$

We will work with this grid for the rest of this paper.

Then we consider the (overlapping) intervals

$$I_k := [x_{k-1}, x_{k+1}], \quad \text{for } k = -(N_\varepsilon - 1), -(N_\varepsilon - 2), \dots, (N_\varepsilon - 1). \quad (4.4)$$

Notice that  $x_{\pm N_\varepsilon}$  is separated from  $x_{\pm(N_\varepsilon-1)}$  by up to length  $2\ell$ , while the rest of the points are separated by length  $\ell$ . Since our energetic estimates will all hold uniformly for subsystems that are sufficiently large, and our large deviation bounds will all hold uniformly for subsystems whose length vary within a compact set, it will not matter that the boundary points may be up to  $2\ell$  away from the neighboring points and we will ignore this issue for the rest of the proof.

### Upper bound.

Here we will prove the upper bound, i.e., that

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \\ \lesssim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 - \gamma}{\varepsilon} \right). \end{aligned}$$

As explained in Section 2, for the upper bound we will work with  $\delta^-$  transition layers, and it will be sufficient to show that for any sufficiently small  $\gamma > 0$  and some sufficiently small  $\delta > 0$ , there is an  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  we have

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \delta^- \text{ transition layers}) \\ \lesssim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 - \gamma}{\varepsilon} \right). \end{aligned} \quad (4.5)$$

Since the probability of transition layers is less than the probability of  $\delta^-$  transition layers, the proof of the upper bound follows immediately.

The subtle part of the proof will be estimating the probability of a transition layer on a subsystem. Recall from Subsection 1.3 that we cannot get the expected cost  $c_0$  by estimating the probability

$$\mu_{\varepsilon, (-2\ell, 2\ell)}^{u_-, u_+} (u \text{ has a } \delta^- \text{ transition layer in } (-\ell, \ell))$$

because of the nontrivial dependence of this probability on the boundary conditions  $u_\pm$ . To avoid this problem, we will use reflection operators to transform  $\delta^-$  transition layers into wasted  $\delta^-$  excursions (see Definition 2.4 and the accompanying discussion).

With this scheme in mind, let us now begin our estimates.



Step 1. Fix  $\gamma > 0$ . Let  $\delta > 0$  be a small constant and  $M < \infty$  be a large constant to be chosen below. Our first step will be to decompose the set of functions in which we are interested. Namely, we notice that the set of continuous paths  $u : [-L_\varepsilon, L_\varepsilon] \rightarrow \mathbb{R}$  satisfying the boundary conditions  $u(\pm L_\varepsilon) = \pm 1$  and exhibiting at least  $(2n+1)\delta^-$  transition layers is contained in the union of the following three sets:

- The set of paths that exhibit an atypically large value at one of the  $x_k$ :

$$\mathcal{A}_1 := \left\{ u : \exists k \in \{-(N_\varepsilon - 1), \dots, (N_\varepsilon - 1)\} : |u(x_k)| \geq M \right\}. \quad (4.6)$$

- The complementary set intersected with the set of paths that are bounded away from  $\pm 1$  on all of  $[x_k, x_{k+1}]$  for some  $k$ :

$$\mathcal{A}_2 := \mathbb{C}\mathcal{A}_1 \cap \left\{ u : \exists k \in \{-(N_\varepsilon - 1), \dots, (N_\varepsilon - 2)\} : \right. \\ \left. u \in [-1 + \delta, 1 - \delta] \text{ on all of } [x_k, x_{k+1}] \right\}. \quad (4.7)$$

- The complement of  $\mathcal{A}_1$  intersected with the set of paths performing  $(2n+1)\delta^-$  transitions, each of which is completely contained in (at least) one of the overlapping intervals  $I_k$ . We denote this set

$$\mathcal{A}_3 := \mathbb{C}\mathcal{A}_1 \cap \left\{ u : \text{there exist } 2n \text{ integers} \right. \\ \left. -N_\varepsilon + 1 \leq k_1 \leq k_2 \leq \dots \leq k_{2n} \leq N_\varepsilon - 1 \text{ such that} \right. \\ \left. \text{in each interval } I_{k_i} \text{ there is a } \delta^- \text{ layer} \right\}.$$

Note that there might be more than one layer in a single interval; the  $2n$ -tuple  $(k_1, \dots, k_{2n})$  allows for a possible higher multiplicity. There may also be *more than*  $2n$  layers; the statement is that there are *at least*  $2n$  layers.

Above, we have made use of the boundary conditions. Indeed, for  $\mathcal{A}_1$ , we have omitted the points  $x_{\pm N_\varepsilon}$  since  $u(\pm L_\varepsilon) = \pm 1$ . For  $\mathcal{A}_2$  we have omitted the boxes at the boundary since the boundary conditions make it impossible that  $u(x) \in [-1 + \delta, 1 - \delta]$  for all  $x$  in the box. For  $\mathcal{A}_3$  we have recalled that the boundary conditions force there to be at least one transition. Even though  $u$  has  $2n+1$  layers, we can expect an additional cost only for the  $2n$  “extra” layers and hence only keep track of  $2n$  layers.

Because the set of interest is contained within the above-mentioned sets, it suffices to bound

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_1) + \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_2) + \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3). \quad (4.8)$$

Step 2. We first give a bound on the probability of  $\mathcal{A}_1$ . This bound follows directly from Lemma 4.1. In fact, we get

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_1) &\leq \sum_{k=-(N_\varepsilon-1)}^{N_\varepsilon-1} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(|u(x_k)| \geq M) \\ &\stackrel{(4.1)}{\leq} \left( 2 \frac{L_\varepsilon}{\ell} - 1 \right) \exp \left( - \frac{M}{\varepsilon C_2} \right) \\ &\leq L_\varepsilon \exp \left( - \frac{M}{\varepsilon C_2} \right). \end{aligned} \quad (4.9)$$

In particular, we can choose  $M$  large enough so that  $M/C_2 \geq 2nc_0$  and

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1) \leq L_\varepsilon \exp\left(-\frac{2nc_0}{\varepsilon}\right).$$

Hence, the probability of  $\mathcal{A}_1$  is of higher order with respect to the right-hand side of (4.5).

We remark that it is here where  $M$  (and therefore also  $\varepsilon_0$ ) acquires a dependence on  $n$ .

Step 3. To bound the second probability in (4.8), we write

$$\begin{aligned} & \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_2) \\ & \leq \sum_{k=-(N_\varepsilon-1)}^{N_\varepsilon-2} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in [-1+\delta, 1-\delta] \text{ on all of } [x_k, x_{k+1}] \right. \\ & \quad \left. \text{and } u(x_{k-1}), u(x_{k+2}) \in [-M, M] \right). \end{aligned} \quad (4.10)$$

Using the Markov property (3.8), we can write for any  $k$

$$\begin{aligned} & \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in [-1+\delta, 1-\delta] \text{ on all of } [x_k, x_{k+1}] \right. \\ & \quad \left. \text{and } u(x_{k-1}), u(x_{k+2}) \in [-M, M] \right) \\ & = \int_{-M}^M \int_{-M}^M \nu_{k-1,k+2}(du_-, du_+) \\ & \quad \times \mu_{\varepsilon,(x_{k-1},x_{k+2})}^{u_-,u_+} \left( u \in [-1+\delta, 1-\delta] \text{ on all of } [x_k, x_{k+1}] \right), \end{aligned} \quad (4.11)$$

where  $\nu_{k-1,k+2}$  denotes the marginal distribution of the pair  $(u(x_{k-1}), u(x_{k+2}))$ . We now want to invoke the large deviation bound (3.17) and the energy bound from Lemma 2.3 for the measures  $\mu_{\varepsilon,(x_{k-1},x_{k+2})}^{u_-,u_+}$ . To this end, we observe that a  $\delta/2$  ball around functions contained in  $[-1+\delta, 1-\delta]$  consists of functions contained in  $[-1+\delta/2, 1-\delta/2]$ . Redefining  $C_1$  by up to a factor of 8 to account for the parameter  $\delta/2$  and interval length (here  $\ell$  rather than  $2\ell$ ), we have that, for any  $\gamma > 0$  and  $\delta > 0$ , there exists an  $\varepsilon_0 > 0$  such that, for all  $\varepsilon \leq \varepsilon_0$  and all  $u_-, u_+ \in [-M, M]$ , there holds

$$\begin{aligned} & \mu_{\varepsilon,(x_{k-1},x_{k+2})}^{u_-,u_+} \left( u(x) \in [-1+\delta, 1-\delta] \text{ on all of } [x_k, x_{k+1}] \right) \\ & \leq \exp\left(-\frac{1}{\varepsilon} \left( \frac{2\delta^2\ell}{C_1} - \gamma \right)\right). \end{aligned} \quad (4.12)$$

(Here we have used the fact that  $L_\varepsilon \gg 1$ , so that we can choose  $\ell > \ell_*$  to satisfy Lemma 2.3.) Letting  $\gamma = 1$  and choosing  $\ell$  so that  $\delta^2\ell \geq C_1$ , the combination of (4.10), (4.11), and (4.12) gives

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_2) \leq L_\varepsilon \exp\left(-\frac{\delta^2\ell}{C_1\varepsilon}\right), \quad (4.13)$$

where we have trivially bounded the integral of  $\nu$  by 1. In particular, for  $\ell$  large but order-one (and depending on  $n, \delta$ ), we have that the probability of  $\mathcal{A}_2$  is also of higher order with respect to the right-hand side of (4.5).

Step 4. Finally, we arrive at the subtler part, in which we will need the reflection operators. To begin with, let  $\bar{k} = (k_1, \dots, k_{2n})$  and write

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3) \leq \sum_{\bar{k} \in \mathcal{I}} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3^{\bar{k}}), \quad (4.14)$$

where  $\mathcal{I}$  is the set of nondecreasing  $2n$ -tuples, i.e.,

$$\mathcal{I} := \left\{ \bar{k} = (k_1, k_2, \dots, k_{2n}) \in \{-(N_\varepsilon - 1), \dots, (N_\varepsilon - 1)\}^{2n} \right. \\ \left. \text{with } k_{i-1} \leq k_i \right\},$$

and

$$\mathcal{A}_3^{\bar{k}} := \mathbb{C}\mathcal{A}_1 \cap \left\{ \text{in each } I_{k_i} \text{ there is a } \delta^- \text{ layer} \right\}. \quad (4.15)$$

The right-hand side of (4.15) is slightly ambiguous if several indices coincide or in the case of overlapping intervals, i.e. if  $k_{i+1} = k_i + 1$  for some  $i$ . If  $j$  subsequent indices coincide, the right-hand side of (4.15) has to be interpreted as saying that there are at least  $j$   $\delta^-$  transitions in the corresponding interval. In the case of overlapping intervals, for instance if  $k_{i+1} = k_i + 1$ , the right-hand side of (4.15) should be interpreted to mean that there are at least two transitions in the interval  $[(k_i - 1)\ell, (k_i + 2)\ell]$  and, moreover, one is fully contained in  $[(k_i - 1)\ell, (k_i + 1)\ell]$  and one is fully contained in  $[k_i\ell, (k_i + 2)\ell]$ .

The index set satisfies

$$|\mathcal{I}| \lesssim N_\varepsilon^{2n} \lesssim (L_\varepsilon)^{2n}. \quad (4.16)$$

(Recall our convention for the use of the symbol  $\lesssim$  introduced in Notation 1.4.) Hence, to complete the proof of (4.5), it suffices to show that for fixed  $\bar{k} \in \mathcal{I}$ , we have

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3^{\bar{k}}) \lesssim \exp\left(-\frac{2nc_0 - \gamma}{\varepsilon}\right). \quad (4.17)$$

As explained above, the main step is to reduce the problem of estimating the probability of  $\delta^-$  layers to estimating the probability of wasted  $\delta^-$  excursions. This will be achieved through suitable *reflections*.

Let us at first assume that the  $I_k$  are well-separated in the sense that

$$k_i \geq k_{i-1} + 4 \quad \text{for all } i.$$

Let us also assume that we are away from the boundary, i.e., that

$$k_1 \geq -N_\varepsilon + 2, \quad k_{2n} \leq N_\varepsilon - 2.$$

We will consider the possibilities of (a) intervals that overlap or are nearby, (b) intervals that are the same ( $k_i = k_{i+1}$ ), and (c) boundary intervals at the end of Step 5.

We start by defining  $n$  left stopping points  $\chi_1, \dots, \chi_n$  in the following manner. For  $i = 1, \dots, n$  we set

$$\chi_i := \inf \left\{ y \geq x_{k_{i-1}} : u(y) = 0 \text{ and } |u(x)| = 1 - \delta \right. \\ \left. \text{for some } x \in (x_{k_{2i-1}-1}, y) \right\}. \quad (4.18)$$

Here we set  $\chi_i = L_\varepsilon$  if the corresponding set is empty. It is easy to see that these random points are all left stopping points. In a similar fashion, for  $i = n+1, \dots, 2n$  we set

$$\chi_i := \sup \left\{ y \leq x_{k_i+1} : u(y) = 0 \text{ and } |u(x)| = 1 - \delta \text{ for some } x \in (y, x_{k_i+1}) \right\}. \quad (4.19)$$

Here we set  $\chi_i = -L_\varepsilon$  if the corresponding set is empty. Then  $\chi_i$  is a right stopping point for all  $i = n+1, \dots, 2n$ . For any  $u \in \mathcal{A}_3^k$ , all the left and right stopping points  $\chi_i$  are contained in the corresponding intervals  $I_{k_i}$  and, furthermore, we have

$$\chi_1 < \chi_2 < \dots < \chi_n < \chi_{n+1} < \dots < \chi_{2n}. \quad (4.20)$$

Finally, note that as soon as  $\chi_i \neq \pm L_\varepsilon$ , we have that  $u(\chi_i) = 0$ .

For any left stopping point  $\chi_l \in \{\chi_1, \dots, \chi_n\}$  and any right stopping point  $\chi_r \in \{\chi_{n+1}, \dots, \chi_{2n}\}$ , we now define the reflection operator  $R_{\chi_l}^{\chi_r}$ . If  $\chi_l < \chi_r$  (which is the case for any  $u \in \mathcal{A}_3^k$  as remarked above), we set

$$R_{\chi_l}^{\chi_r} u(x) := \begin{cases} -u(x) & \text{for } x \in [\chi_l, \chi_r], \\ u(x) & \text{for } x \notin [\chi_l, \chi_r]. \end{cases}$$

If  $\chi_l \geq \chi_r$  we set  $R_{\chi_l}^{\chi_r} u := u$ . We clearly have  $R_{\chi_l}^{\chi_r} R_{\chi_l}^{\chi_r} = \text{Id}$ ; hence,  $R_{\chi_l}^{\chi_r}$  is injective and onto. In order to show that  $R_{\chi_l}^{\chi_r}$  preserves  $\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}$ , we observe that for any measurable and bounded test function  $\Phi: C([-L_\varepsilon, L_\varepsilon]) \rightarrow \mathbb{R}$ , we have

$$\begin{aligned} & \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\Phi \circ R_{\chi_l}^{\chi_r}) \\ &= \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l < \chi_r\}} \Phi \circ R_{\chi_l}^{\chi_r}) + \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l \geq \chi_r\}} \Phi \circ R_{\chi_l}^{\chi_r}) \\ &= \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} \left( \mathbf{1}_{\{\chi_l < \chi_r\}} \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\Phi \circ R_{\chi_l}^{\chi_r} | \mathcal{F}_{[-L_\varepsilon, \chi_l]} \vee \mathcal{F}_{[\chi_r, L_\varepsilon]}) \right) \\ & \quad + \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l \geq \chi_r\}} \Phi) \\ &\stackrel{(3.10)}{=} \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} \left( \mathbf{1}_{\{\chi_l < \chi_r\}} \mathbb{E}_{(\chi_l, \chi_r)}^{\mu_\varepsilon, \mathbf{u}} (\Phi \circ R_{\chi_l}^{\chi_r}) \right) \\ & \quad + \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l \geq \chi_r\}} \Phi). \end{aligned}$$

Now we can use the fact that on the set  $\{\chi_l < \chi_r\}$  we have almost surely that  $u(\chi_l) = u(\chi_r) = 0$  and the invariance of the measure  $\mu_{\varepsilon, (\chi_l, \chi_r)}^{0,0}$  under the reflection  $R: u \mapsto -u$ . Note that the latter property relies on the symmetry of the double-well potential  $V$ . We get

$$\begin{aligned} & \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} \left( \mathbf{1}_{\{\chi_l < \chi_r\}} \mathbb{E}_{(\chi_l, \chi_r)}^{\mu_\varepsilon, \mathbf{u}} (\Phi \circ R_{\chi_l}^{\chi_r}) \right) + \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l \geq \chi_r\}} \Phi) \\ &= \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} \left( \mathbf{1}_{\{\chi_l < \chi_r\}} \mathbb{E}_{(\chi_l, \chi_r)}^{\mu_\varepsilon, \mathbf{u}} (\Phi) \right) + \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\mathbf{1}_{\{\chi_l \geq \chi_r\}} \Phi) \\ &\stackrel{(3.10)}{=} \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1, 1} (\Phi). \end{aligned} \quad (4.21)$$

Now we are finally ready to define the reflection operator as the composition

$$R := R_{\chi_1}^{\chi_{2n}} \circ \dots \circ R_{\chi_{n-1}}^{\chi_{n+2}} \circ R_{\chi_n}^{\chi_{n+1}}. \quad (4.22)$$

We have again that  $R^2 = \text{Id}$ . For any profile  $u \in \mathcal{A}_3^k$ , the operator  $R$  acts in the following way: In intervals of the form  $(\chi_i, \chi_{i+1})$  for  $i$  odd,  $u$  is replaced by  $-u$ ,

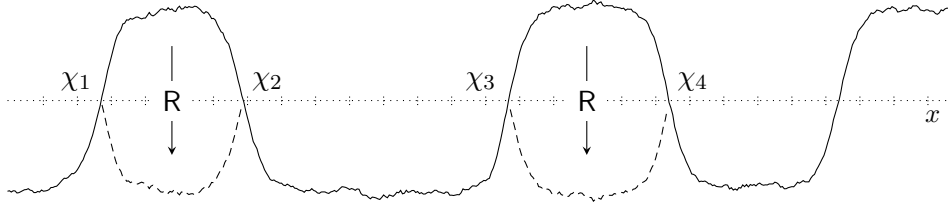


FIGURE 4.1. A typical path in  $\mathcal{A}_3$ . The reflection operator  $R$  turns the up and down transitions in the intervals  $I_{k_i}$  into wasted excursions in the same intervals.

and on the rest of the system,  $u$  is left invariant. The action of the operator  $R$  on a typical path in  $\mathcal{A}_3$  is illustrated in Figure 4.1.

Finally, define the reflection of a set  $\mathcal{A}$  as

$$R\mathcal{A} = \{v : v = Ru \text{ for some } u \in \mathcal{A}\}.$$

As a composition of measure-preserving transformations, the operator  $R$  preserves  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  as well. Hence, we have in particular that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_3^{\bar{k}}) = \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(R\mathcal{A}_3^{\bar{k}}). \quad (4.23)$$

This is useful because for  $u \in \mathcal{A}_3^{\bar{k}}$  the profile  $Ru$  has a wasted  $\delta^-$  excursion on each interval  $I_{k_i}$  (as is easy to check). In other words, we note that  $R\mathcal{A}_3^{\bar{k}}$  is a (proper) subset of the functions with wasted  $\delta^-$  excursions in the given intervals.

Step 5. It remains to bound the probability of the sets  $R\mathcal{A}_3^{\bar{k}}$ . Again, let us at first assume that the  $I_k$  are well-separated and away from the boundary in the sense described above. We consider the more general case at the end of this step.

Using the Markov property again, we have

$$\begin{aligned} & \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(R\mathcal{A}_3^{\bar{k}}) \\ & \leq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathcal{CA}_1 \text{ and in each } I_{k_i} \text{ there is a wasted } \delta^- \text{ excursion}) \\ & = \int_{-M}^M \cdots \int_{-M}^M \nu_{k_1-2,k_1+2,k_2-2,\dots,k_{2n}+2}(du_{k_1-2}, du_{k_1+2}, du_{k_2-2}, \dots, du_{k_{2n}+2}) \\ & \quad \prod_{i=1}^{2n} \mu_{\varepsilon,(x_{k_i-2},x_{k_i+2})}^{u_{k_i-2},u_{k_i+2}} \left( \text{there is a wasted } \delta^- \text{ excursion in } I_{k_i} \right), \end{aligned} \quad (4.24)$$

where  $\nu_{k_1-2,k_1+2,k_2-2,\dots,k_{2n}+2}$  denotes the distribution of the  $4n$ -dimensional marginal  $u(x_{k_1-2}), u(x_{k_1+2}), u(x_{k_2-2}), \dots, u(x_{k_{2n}+2})$ .

Now we would like to apply the large deviation bound (3.17) and the energy bound from Lemma 2.5. We observe that a  $\delta$  ball around paths with a wasted  $\delta^-$  excursion is equal to the set of paths with a wasted  $(2\delta)^-$  excursion. As a result, we get that for any  $\gamma > 0$  and  $\delta > 0$  there exists an  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  and for all boundary data contained in  $[-M, M]$ , the probability of a wasted  $\delta^-$

excursion is bounded by

$$\begin{aligned} \mu_{\varepsilon, (x_{k_i-2}, x_{k_i+2})}^{u_{k_i-2}, u_{k_i+2}} & \left( \text{there is a wasted } \delta^- \text{ excursion in } I_{k_i} \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} (c_0 - 2C\delta - \gamma) \right). \end{aligned} \quad (4.25)$$

Choosing  $\delta$  sufficiently small with respect to  $\gamma$  and estimating the integral of  $\nu$  by 1 as usual, we have from the combination of (4.23), (4.24), and (4.25) that (4.17) holds (up to a redefinition of  $\gamma$ ). Thus, finally, (4.14), (4.16), and (4.17) imply

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3) \lesssim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 - \gamma}{\varepsilon} \right),$$

which concludes the proof of the upper bound in the well-separated case.

It remains to consider the three special cases: (a) intervals that overlap or are nearby, (b) intervals that are the same ( $k_i = k_{i+1}$ ), (c) intervals that are boundary intervals.

**Case (a)** If two or more intervals overlap (i.e., if  $k_i = k_{i-1} + 1$ ) or are nearby (i.e., if  $k_{i-1} + 2 \leq k_i \leq k_{i-1} + 3$ ), then we lump them together into a single, larger interval and proceed as in (b), below. The size of the largest possible interval formed in this way is  $(4 + 3(2n - 1))\ell$ . Our energy estimates require only that the interval length be sufficiently large and our large deviation estimates are uniform as long as the interval length falls within a compact set. (Here we rely on the fact that  $n$  is order-one with respect to  $\varepsilon$ .)

**Case (b)** If a multi-index  $\bar{k}$  has repeated indices so that there is more than one  $\delta^-$  transition layer in a single interval, then we will use large deviation estimates for the event of having *more than one wasted  $\delta^-$  excursion* in a single interval.

Assume that we have  $k_j = k_{j+1} = \dots k_{j+m}$  for some  $j < 2n$  and some  $1 \leq m \leq 2n$ . Furthermore, assume that the set of  $m + 1$  indices is maximal in the sense that either  $j = 1$  or  $k_{j-1} \leq k_j - 4$  and similarly that either  $j + m = 2n$  or  $k_{j+m+1} \geq k_{j+m} + 4$ . In this case, we define the  $m + 1$  stopping points  $\chi_j, \dots \chi_{j+m}$  in the following way.

Consider any index  $i \in \{j, \dots, j + m\}$  that satisfies  $i \leq n$ . For  $i = j$ , we define  $\chi_j$  as in (4.18). On the other hand, for  $i > j$ , we define

$$\chi_i := \inf \left\{ y \geq x_{k_j-1} : u(y) = 0 \text{ and there are } (i - j) \delta^- \text{ layers in } (x_{k_j-1}, y) \right\}. \quad (4.26)$$

As usual, we define  $\chi_i = L_\varepsilon$  if the set above is empty.

Now consider any index  $i \in \{j, \dots, j + m\}$  that satisfies  $i > n$ . For  $i = j + m$ , we define  $\chi_{j+m}$  as in (4.19). On the other hand, for  $i < j + m$ , we define

$$\begin{aligned} \chi_i &:= \sup \left\{ y \leq x_{k_{j+1}} : u(y) = 0 \right. \\ &\quad \left. \text{and there are } (m - (i - j)) \delta^- \text{ layers in } (y, x_{k_{j+1}}) \right\}. \end{aligned} \quad (4.27)$$

Again, we take the usual definition  $\chi_i = -L_\varepsilon$  if the set above is empty.

As above these random points  $\chi_i$  are left stopping points for  $i \leq n$  and right stopping points for  $i \geq n + 1$ . Furthermore, we still have that (4.20) holds for all  $u \in \mathcal{A}_3^k$ . The measure preserving reflection operator  $R$  can be defined as above in (4.22), and  $R$  maps each  $u \in \mathcal{A}_3^{\bar{k}}$  to a path that has  $m + 1$  wasted  $\delta^-$  excursions in  $I_{k_j}$ . (Specifically, we mean  $m + 1$  wasted  $\delta^-$  excursions on intervals

$(x_-^i, x_+^i) \subset I_{k_j}$  for  $i \in \{j, \dots, j+m\}$  that are mutually disjoint except for possibly the endpoints.)

We leave it to the reader to verify that a generalization of Lemma 2.5 is:

**Lemma 4.2.** *There exists  $C < \infty$  with the following property. Fix  $\delta > 0$ . For any system sizes  $\ell_1, \ell_2 < \infty$  sufficiently large and boundary conditions  $u_{\pm} \in \mathbb{R}$ , set*

$$\begin{aligned} \mathcal{A}^{\text{bc}} &:= \{u \in C([- \ell_1 - \ell_2, \ell_1 + \ell_2]): u(-\ell_1 - \ell_2) = u_-, u(\ell_1 + \ell_2) = u_+\}, \\ \mathcal{A}_0^{\text{bc}} &:= \{u \in \mathcal{A}^{\text{bc}}: u \text{ has } m \text{ disjoint wasted } \delta^- \text{ excursions in } (-\ell_1, \ell_1)\}. \end{aligned}$$

Define the optimal cost

$$c_\ell := \frac{1}{m} \left( \inf_{\mathcal{A}_0^{\text{bc}}} E_{(-\ell_1 - \ell_2, \ell_1 + \ell_2)}(u) - \inf_{\mathcal{A}^{\text{bc}}} E_{(-\ell_1 - \ell_2, \ell_1 + \ell_2)}(u) \right).$$

Then uniformly with respect to the boundary values  $u_{\pm}$ , one has

$$c_\ell - c_0 \leq C \delta.$$

**Case (c)** Suppose for instance that there is a transition layer in  $(x_{-N_\varepsilon}, x_{-N_\varepsilon+1})$ . Then we know the boundary value  $u(x_{-N_\varepsilon}) = u(-L_\varepsilon) = -1$ , while the boundary value  $u(x_{N_\varepsilon-2})$  at the other end of the subinterval is unknown. This is easily handled by a suitable “one-sided” generalization of Lemma 2.5, which is easy to prove.

Using the facts from above, the proof of the upper bound is completed by decomposing  $\mathcal{A}_3^k$  into the various cases and recovering the correct (and identical) bounds in each case.

### Lower bound.

We turn now to the matching lower bound, i.e., that

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n+1) \text{ transition layers}) \\ \gtrsim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 + \gamma}{\varepsilon} \right). \end{aligned}$$

As explained in Subsection 2, for the lower bound we will work with  $\delta^+$  transition layers (cf. Definition 2.6). Because of the boundary conditions and the definition of  $\delta^+$  layers, it will be sufficient to show that, for some  $\delta \in (0, 1/2)$ , we have

$$\begin{aligned} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (u \text{ has } (2n) \delta^+ \text{ transition layers}) \\ \gtrsim (L_\varepsilon)^{2n} \exp \left( -\frac{2nc_0 + \gamma}{\varepsilon} \right). \end{aligned} \tag{4.28}$$

Indeed, in analogy with the upper bound, the probability of  $\delta^+$  layers is bounded above by the probability of transition layers, and because of the boundary conditions there must be an odd number of transitions.

Step 1. Once again, we will use the gridpoints  $x_k$  defined in (4.3). Our first step is to get some control on the values of  $u$  at the gridpoints. The following lemma, used below, is established via techniques similar to those used for the upper bound.

**Lemma 4.3.** *For any  $M < \infty$  sufficiently large, there exists  $\ell_* < \infty$  and  $\varepsilon_0 > 0$  such that, for  $\ell \geq \ell_*$  and  $\varepsilon \leq \varepsilon_0$ , we have for any  $L_\varepsilon$  satisfying (1.10) that*

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} \left( u \in \mathcal{A}_1 : u(x) \leq 0 \text{ for all } x \in [-L_\varepsilon, -2\ell] \right) \geq \frac{1}{3}. \quad (4.29)$$

Recall the definition of  $\mathcal{A}_1$  in (4.6).

The proof is similar to the proof of the upper bound, and is deferred to Subsection 6.5. The main idea is that while the boundary conditions force there to be a transition layer, with high probability, there is *only one transition layer*. Moreover, by symmetry, this layer is as likely to appear on  $[0, L_\varepsilon]$  as it is on  $[-L_\varepsilon, 0]$  (hence neither probability can be more than  $1/2$ ). On the other hand, for  $u$  to hit zero away from the transition layer is energetically unlikely, by arguments similar to those used for the upper bound.

Step 2. With Lemma 4.3 in hand, we turn to the basic set-up for the lower bound. In this case, we will not want to use overlapping subintervals. We will also not work with the full system, but only with intervals on the left-hand side. Specifically, we will work with

$$I_k = [x_{k-1}, x_{k+1}] \quad \text{for } k \in \{-(N_\varepsilon - 4), -(N_\varepsilon - 8), \dots, -4\} =: E.$$

We have assumed without loss of generality that 4 divides  $N_\varepsilon$ . (If not, then  $N_\varepsilon = 4j + r$  for some  $j \in \mathbb{N}$  and  $r \in \{1, 2, 3\}$ . Replace  $N_\varepsilon$  by  $N_\varepsilon - r$  throughout.)

We remark that, as usual, for an event falling in the interval  $I_k$ , we will condition on the boundary values on a larger interval. Specifically, we will use a Markov decomposition in which we condition on the boundary values of the enlarged interval

$$\tilde{I}_k := [x_{k-2}, x_{k+2}].$$

Notice that for all  $k \in E$ , the enlarged intervals  $\tilde{I}_k$  are nonintersecting. For future reference, let us denote the set of boundary indices

$$E_b := \{-(N_\varepsilon - 2), -(N_\varepsilon - 6), \dots, -2\}.$$

The rough idea is to consider sets of functions having  $2n$  layers with a layer in one of the intervals  $I_k$  for  $2n$  distinct values of  $k \in E$ . Unfortunately, because we work with functions  $u$  that have *at least*  $2n + 1$  transitions rather than *exactly*  $2n + 1$  transitions, a given function  $u$  may have more than  $2n + 1$  layers and belong to more than one of the sets we have just described. Hence we cannot translate the probability of the union into the sum of the probabilities. In order to work around this, we will work with more restrictive sets.

Analogous to the set  $\mathcal{A}_1$  defined in (4.6) above, we define the following set. Rather than keeping track of all the boundary values, it will be convenient to track only the boundary values for the extended intervals described above. That is, we consider

$$\tilde{\mathcal{A}}_1 := \left\{ u : |u(x_k)| \geq M \text{ for some } k \in E_b \right\}.$$

We now introduce a set that is analogous to the set  $\mathcal{A}_3^{\bar{k}}$  above (but more restrictive, for the reason we have explained). For ease of notation, we do not introduce



a new label. Let  $\bar{k} = (k_1, \dots, k_{2n})$  and consider the set

$$\begin{aligned} \mathcal{A}_3^{\bar{k}} := & \mathbb{C}\tilde{\mathcal{A}}_1 \cap \left\{ \begin{array}{l} \text{in each } I_{k_i} \text{ with } i \text{ odd there exists a } \delta^+ \text{ up layer} \\ \text{and in each } I_{k_i} \text{ with } i \text{ even there exists a } \delta^+ \text{ down layer and} \\ \text{for } k \in E \setminus \{k_i, 1 \leq i \leq 2n\} \text{ } u \text{ does not have a } \delta^+ \text{ layer in } I_k \end{array} \right\}. \end{aligned}$$

Clearly, we have the following inclusion of sets of paths:

$$\{ u \text{ has } (2n) \delta^+ \text{ transition layers} \} \supseteq \bigcup_{\bar{k} \in \mathcal{I}} \mathcal{A}_3^{\bar{k}}, \quad (4.30)$$

where  $\mathcal{I}$  is the following set of *well-separated indices on the negative x-axis*:

$$\mathcal{I} := \left\{ \bar{k} = (k_1, k_2, \dots, k_{2n}) \in E^{2n} : \text{for all } i, k_{i-1} < k_i \right\}.$$

Moreover, the sets  $\mathcal{A}_3^{\bar{k}}$  for  $\bar{k} \in \mathcal{I}$  are disjoint, so that (4.30) implies

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} ( u \text{ has } (2n) \delta^+ \text{ transition layers} ) \geq \sum_{\bar{k} \in \mathcal{I}} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (\mathcal{A}_3^{\bar{k}}). \quad (4.31)$$

The set on the right-hand side of (4.30) is certainly smaller than the set on the left-hand side, but the bound will be good enough on the level of scaling since

$$|\mathcal{I}| \gtrsim N_\varepsilon^{2n} \gtrsim (L_\varepsilon)^{2n}. \quad (4.32)$$

Step 3. Given (4.31) and (4.32), we will be done if we can establish that for any  $\gamma > 0$  and for  $\varepsilon > 0$  sufficiently small, we have

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} (\mathcal{A}_3^{\bar{k}}) \gtrsim \exp \left( - \frac{2nc_0 + \gamma}{\varepsilon} \right). \quad (4.33)$$

To this end, fix any multi-index  $\bar{k} \in \mathcal{I}$ . We will now bound the probability of  $\mathcal{A}_3^{\bar{k}}$  using reflections, as we did for the upper bound. Indeed, let

$$\begin{aligned} \chi_{2i-1} := & \inf \left\{ y \in I_{k_{2i-1}} : u(y) = 0, \quad u(x) = -1 - \delta \right. \\ & \left. \text{for some } x \in (x_{k_{(2i-1)-1}}, y) \right\}, \\ \chi_{2i} := & \sup \left\{ y \in I_{k_{2i}} : u(y) = 0, \quad u(x) = -1 - \delta \right. \\ & \left. \text{for some } x \in (y, x_{k_{2i}+1}) \right\}. \end{aligned}$$

Then we define the reflection operator  $R$  as

$$R = R_{\chi_{2n-1}}^{\chi_{2n}} \circ \dots \circ R_{\chi_1}^{\chi_2}.$$

By the same argument as above in (4.21) it can be seen that this operator preserves the measure  $\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}$ . Notice that  $R$  creates  $\delta^+$  wasted excursions in the intervals  $I_{k_i}$  and cannot create layers in any interval  $I_k$  for  $k \in E \setminus \{k_i, 1 \leq i \leq 2n\}$ . We

recover

$$\begin{aligned}
& \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3^{\bar{k}}) \\
&= \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{R}\mathcal{A}_3^{\bar{k}}) \\
&= \int_{-M}^M \cdots \int_{-M}^M \nu_{-(N_\varepsilon-2), -(N_\varepsilon-6), \dots, -2}(du_{-(N_\varepsilon-2)}, du_{-(N_\varepsilon-6)}, \dots, du_{-2}) \\
&\quad \prod_{i=1}^{2n} \mu_{\varepsilon, (x_{k_i-2}, x_{k_i+2})}^{u_{k_i-2}, u_{k_i+2}} \left( \text{there is a wasted } \delta^+ \text{ excursion in } I_{k_i} \right) \\
&\quad \prod_{k \in E \setminus \{k_i, 1 \leq i \leq 2n\}} \mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_{k-2}, u_{k+2}} \left( \text{there is no } \delta^+ \text{ layer in } I_k \right) \\
&\geq \int_{-M}^0 \cdots \int_{-M}^0 \nu_{-(N_\varepsilon-2), -(N_\varepsilon-6), \dots, -2}(du_{-(N_\varepsilon-2)}, du_{-(N_\varepsilon-6)}, \dots, du_{-2}) \\
&\quad \prod_{i=1}^{2n} \mu_{\varepsilon, (x_{k_i-2}, x_{k_i+2})}^{u_{k_i-2}, u_{k_i+2}} \left( \text{there is a wasted } \delta^+ \text{ excursion in } I_{k_i} \right) \\
&\quad \prod_{k \in E \setminus \{k_i, 1 \leq i \leq 2n\}} \mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_{k-2}, u_{k+2}} \left( \text{there is no } \delta^+ \text{ layer in } I_k \right). \tag{4.34}
\end{aligned}$$

As usual,  $\nu$  denotes the distribution of boundary values, here at the boundary points of each extended interval  $\tilde{I}_k$  for  $k \in E$ . Note that the second equality follows from the definition of wasted  $\delta^+$  excursions. The definition of wasted  $\delta^-$  excursions is different and led to an inequality in the analogous estimate, cf. (4.24).

We remark that we do not actually need to condition on the boundary values for every  $k \in E$ —it would be enough to consider the intervals  $\tilde{I}_k$  for  $k \in \bar{k}$  and the complementary intervals—but doing it this way keeps notation simple and because of (1.10), it does not affect our bound by more than an exponentially small amount.

We now turn to the lower large deviation bound (3.18) and the energy bound from Lemma 2.8 (where we use that the boundary values are in  $[-M, 0]$ ). We recall that the set  $\mathcal{A}_{\delta, pre}^{bc}$  from (2.7) was defined precisely so that

$$B(\mathcal{A}_{\delta, pre}^{bc}, \delta) = \{u : u \text{ has a wasted } \delta^+ \text{ excursion in } [-\ell, \ell]\}.$$

Therefore, applying the large deviation estimate to (4.34), we conclude that for any  $\gamma > 0$  and  $\delta > 0$  small enough, there exists an  $\varepsilon_0 > 0$  such that for any  $\bar{k} \in \mathcal{I}$  and any  $\varepsilon \leq \varepsilon_0$ , we have

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_3^{\bar{k}}) \geq \exp\left(-\frac{2nc_0 + \gamma}{\varepsilon}\right) \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}(\mathcal{A}_4^{\bar{k}}), \tag{4.35}$$

where

$$\begin{aligned}
\mathcal{A}_4^{\bar{k}} := & \left\{ u : u(x_k) \in [-M, 0] \text{ for all } k \in E_b \text{ and} \right. \\
& \left. u \text{ has no layer in } I_k \text{ for any } k \in E \setminus \{k_i, 1 \leq i \leq 2n\} \right\}.
\end{aligned}$$

At the same time, for any  $\bar{k} \in \mathcal{I}$  we have

$$\mathcal{A}_4^{\bar{k}} \supseteq \{u \in \mathbb{C}\mathcal{A}_1 : u(x) \leq 0 \text{ for all } x \in [-L_\varepsilon, -2\ell]\},$$

where  $\mathcal{A}_1$  includes all the gridpoints, as defined in (4.6). Hence by the estimate (4.29) from Lemma 4.3, the lower bound (4.35) improves to

$$\mu_{\varepsilon, (-L\varepsilon, L\varepsilon)}^{-1,1}(\mathcal{A}_3^k) \geq \frac{1}{3} \exp\left(-\frac{2nc_0 + \gamma}{\varepsilon}\right), \quad (4.36)$$

which establishes (4.33) and completes the proof of the lower bound.  $\square$

## 5. PROOF OF THEOREM 1.9: THE UNIFORM DISTRIBUTION OF THE LAYER LOCATION

As pointed out in Subsection 1.3, the proof of Theorem 1.9 relies on the construction of a measure-preserving operator  $R_{y,z}$ . This operator maps paths that exhibit a transition near  $y$  to paths that exhibit a transition near  $z$ . It is constructed by performing a point reflection between hitting points of  $\pm 1$  near  $y$  and  $z$ .

The main difficulty of the proof is to show that these hitting points exist with very high probability on the set of paths that perform a transition near  $y$ . The argument for this is provided in the following two lemmas.

The first lemma states, roughly speaking, that in the “bulk,” fluctuations around  $\pm 1$  are of order  $\varepsilon^{1/2}$ . The system needs  $O(|\log \varepsilon|)$  space to relax to this scale. For simplicity, we state the lemma for paths that stay close to 1. By symmetry, the analogous statement holds near  $-1$ .

**Lemma 5.1.** *There exists  $C < \infty$  with the following property. For every  $\ell_0 < \infty$  sufficiently large, there exists  $\varepsilon'_0 > 0$  such that the following holds. For every  $\varepsilon$  and  $\varepsilon_0$  with  $\varepsilon \leq \varepsilon_0 \leq \varepsilon'_0$ , there exists  $K_\varepsilon \in \mathbb{N}$  with*

$$K_\varepsilon \sim \log\left(\sqrt{\frac{\varepsilon_0}{\varepsilon}}\right)$$

such that for

$$\ell_\varepsilon := (2K_\varepsilon + 1)\ell_0$$

and all  $u_\pm \in [1/2, 3/2]$ , we have

$$\begin{aligned} & \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-\ell_0, \ell_0]} |u(x) - 1| \geq \sqrt{\frac{\varepsilon}{\varepsilon_0}} \mid \right. \\ & \quad \left. |u(\pm(2k-1)\ell_0) - 1| \leq \frac{1}{2}, k = 1, 2, \dots, K_\varepsilon \right) \\ & \leq 4 \exp\left(-\frac{1}{C\varepsilon_0}\right). \end{aligned}$$

We present the proof in Subsection 6.6. Next we need a lemma that says that with positive probability, the path *actually hits*  $\pm 1$ . Again, we state the result for hitting points of  $+1$ . By symmetry, the analogous statement holds for hitting points of  $-1$ .

**Lemma 5.2.** *For any  $\ell_0 < \infty$  sufficiently large, there exist  $\varepsilon_0 > 0$  and  $\lambda \in (0, 1)$  such that the following holds. For any  $u_\pm \in [1/2, 3/2]$ , any  $\varepsilon \leq \varepsilon_0$ , and  $K_\varepsilon, \ell_\varepsilon$  as*

in Lemma 5.1, we get

$$\mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \exists x \in [-\ell_0, \ell_0] \text{ such that } u(x) = 1 \right. \\ \left. \left| u(\pm(2k-1)\ell_0) - 1 \right| \leq \frac{1}{2}, k = 1, 2, \dots, K_\varepsilon \right) \geq 1 - \lambda.$$

The proof of this lemma, also presented in Subsection 6.6, follows as a corollary to the previous result by a rescaling argument.

*Proof of Theorem 1.9.* We will show that for some  $\delta \in (0, 1/2)$  and any  $\alpha > 0$ , we have for  $\varepsilon$  sufficiently small that

$$1 - \alpha \leq \frac{L_\varepsilon}{d_\varepsilon} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1, 1} (\text{at least one } \delta^- \text{ up layer of length } \leq 2\ell \\ \text{in } [y - d_\varepsilon, y + d_\varepsilon]) \leq 1 + \alpha. \quad (5.1)$$

At the end of the proof, it will not be hard to improve from a  $\delta^-$  up layer of length less than or equal to  $2\ell$  to a full up transition layer.

**Notation 5.3.** For brevity, we will often say “a transition layer  $\leq 2\ell$ ” as shorthand for “a transition layer of length less than or equal to  $2\ell$ .”

For  $\varepsilon$  small enough we consider intervals of type  $J_{y, \varepsilon} = [y - d_\varepsilon, y + d_\varepsilon] \subseteq [L_\varepsilon, L_\varepsilon]$ . The main step of our argument consists of proving that the probabilities of transitions in these intervals  $J_{y, \varepsilon}$  for different values of  $y$  are roughly the same. Hence fix two points  $y, z$  such that  $J_{y, \varepsilon}, J_{z, \varepsilon} \subseteq [-L_\varepsilon, L_\varepsilon]$ . Without loss of generality, assume that  $y \leq z$ .

As above in the proof of Theorem 1.5, let  $\ell$  and  $M$  be a large constants to be fixed later, and let  $N_\varepsilon$  and  $x_{\pm k}$  be as defined in (4.2) and (4.3). Moreover, consider the overlapping intervals  $I_k = [x_{k-1}, x_{k+1}]$  as in (4.4). Finally, define as in (4.6) the “bad set”  $\mathcal{A}_1$  of functions that have boundary values larger than  $M$  in magnitude. In (4.9) above, we have already established that there is a universal constant  $C_2 < \infty$  such that

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1, 1}(\mathcal{A}_1) \leq L_\varepsilon \exp\left(-\frac{M}{C_2 \varepsilon}\right).$$

Hence, for the system sizes  $L_\varepsilon$  that we consider, the probability of  $\mathcal{A}_1$  can be made arbitrarily small by choosing  $M$  large.

We now define the set of functions

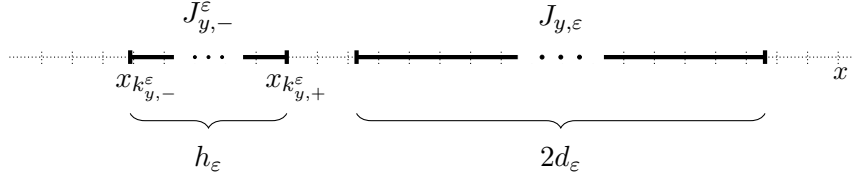
$$\mathcal{J}_{y, \varepsilon} := \{u \in \mathbb{C}\mathcal{A}_1 : u \text{ has a } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y, \varepsilon}\}. \quad (5.2)$$

The set  $\mathcal{J}_{z, \varepsilon}$  is defined analogously. In Steps 1–3 below, we will establish that the probabilities of the  $\mathcal{J}_{y, \varepsilon}$  and  $\mathcal{J}_{z, \varepsilon}$  are roughly the same. The bounds that we obtain will be uniform with respect to  $y$  and  $z$ . Finally, in Step 4 we will show how this implies (5.1), and in Step 5 we will improve to the statement of Theorem 1.9.

**Step 1.** The first step consists of proving that on the set  $\mathcal{J}_{y, \varepsilon}$ , with high probability, the profile  $u$  is close to  $-1$  on a sufficiently large interval  $J_{y, -}^\varepsilon$  just to the left of  $J_{y, \varepsilon}$  and close to  $+1$  on a sufficiently large interval  $J_{z, +}^\varepsilon$  just to the right of  $J_{z, \varepsilon}$ .

The length  $h_\varepsilon$  of each of these auxiliary intervals  $\mathcal{J}_{y, -}^\varepsilon$  and  $\mathcal{J}_{z, +}^\varepsilon$  will be chosen below such that

$$|\log \varepsilon| \ll h_\varepsilon \ll d_\varepsilon.$$

FIGURE 5.1. The interval  $J_{y,\varepsilon}$  and the auxiliary interval  $J_{y,-}^\varepsilon$  to its left.

We first fix the “inner” boundary points of  $J_{y,-}^\varepsilon$  and  $J_{z,+}^\varepsilon$ : In units of  $\ell$ , we set

$$\begin{aligned} k_{y,+}^\varepsilon &:= \max \{k : x_k \leq y - d_\varepsilon\} - 2, \\ k_{z,-}^\varepsilon &:= \min \{k : x_k \geq z + d_\varepsilon\} + 2. \end{aligned}$$

Let  $K_\varepsilon$  be as in the statement of Lemma 5.1. The idea is to make the probability of hitting  $\pm 1$  on the auxiliary intervals large by concatenating many subintervals of length  $K_\varepsilon \ell$  and applying Lemma 5.2 on each subinterval. With this end in mind, we fix integers  $\bar{K}_\varepsilon$  such that

$$\bar{K}_\varepsilon \gg 1 \quad \text{and} \quad h_\varepsilon := \ell(2K_\varepsilon + 1)\bar{K}_\varepsilon \ll d_\varepsilon. \quad (5.3)$$

Then we set

$$k_{y,-}^\varepsilon := k_{y,+}^\varepsilon - (2K_\varepsilon + 1)\bar{K}_\varepsilon \quad \text{and} \quad k_{z,+}^\varepsilon := k_{z,-}^\varepsilon + (2K_\varepsilon + 1)\bar{K}_\varepsilon,$$

and finally

$$J_{y,-}^\varepsilon := [k_{y,-}^\varepsilon \ell, k_{y,+}^\varepsilon \ell] \quad \text{and} \quad J_{z,+}^\varepsilon := [k_{z,-}^\varepsilon \ell, k_{z,+}^\varepsilon \ell].$$

(See Figure 5.1 for an illustration of  $J_{y,\varepsilon}$  and  $J_{y,-}^\varepsilon$ .)

We also define the following sets of indices

$$\mathcal{I}_-^\varepsilon := \{k : k_{y,-}^\varepsilon \leq k \leq k_{y,+}^\varepsilon\}, \quad \mathcal{I}_+^\varepsilon := \{k : k_{z,-}^\varepsilon \leq k \leq k_{z,+}^\varepsilon\}.$$

For later use in (6.81) in the proof of Lemma 5.5, we will make the additional growth assumption

$$|\mathcal{I}_\pm^\varepsilon| = (2K_\varepsilon + 1)\bar{K}_\varepsilon + 1 \leq \exp(c_1/4\varepsilon), \quad (5.4)$$

where  $c_1 > 0$  is defined in (6.86), below. This is not a strong condition; we will typically think of  $h_\varepsilon$  as being much smaller.

Finally, we define another set of “unlikely” paths, paths that have extra  $\delta^-$  layers to the left of  $J_{y,\varepsilon}$  or to the right of  $J_{z,\varepsilon}$ :

$$\begin{aligned} \mathcal{A}_{y,3}^- &:= \{u \in \mathcal{J}_{y,\varepsilon} : \text{there exists } x \leq (k_{y,+}^\varepsilon + 1)\ell \text{ with } u(x) \geq 1 - \delta\}, \\ \mathcal{A}_{y,3}^+ &:= \{u \in \mathcal{J}_{y,\varepsilon} : \text{there exists } x \geq (k_{z,-}^\varepsilon - 1)\ell \text{ with } u(x) \leq -1 + \delta\}, \\ \mathcal{A}_{y,3} &:= \mathcal{A}_{y,3}^- \cup \mathcal{A}_{y,3}^+. \end{aligned} \quad (5.5)$$

We now introduce two lemmas. The proofs of both lemmas are given in Subsection 6.6. The first lemma is an extension of the upper bound in Theorem 1.5 and states roughly speaking that *conditioned* on having a transition in a given interval, the probability of extra layers somewhere else is small.

**Lemma 5.4.** *Let  $Y$  be a subinterval of  $[-L_\varepsilon, L_\varepsilon]$  and let  $x_- = k_- \ell$  and  $x_+ = k_+ \ell$  be two gridpoints (cf. (4.3)) to the left and to the right of  $Y$  respectively with distance  $\geq \ell$  from  $Y$ . We denote by  $\mathcal{J}_Y$  and  $\mathcal{A}_{Y,3}$  the sets*

$$\begin{aligned}\mathcal{J}_Y &:= \{u \in \mathbb{C}\mathcal{A}_1 : u \text{ has a } \delta^- \text{ up layer in } Y\}, \\ \mathcal{A}_{Y,3} &:= \{u \in \mathcal{J}_Y \text{ and } u \text{ has another } \delta^- \text{ layer outside of } [x_-, x_+]\}.\end{aligned}$$

Fix any  $\gamma > 0$  and any  $M < \infty$  sufficiently large. For any  $\delta > 0$  sufficiently small and  $\ell < \infty$  sufficiently large, there exists  $\varepsilon_0 > 0$  such that, for all  $\varepsilon \leq \varepsilon_0$ , we have

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_{Y,3}) \lesssim L_\varepsilon \exp\left(-\frac{c_0 - \gamma}{\varepsilon}\right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_Y). \quad (5.6)$$

We now apply Lemma 5.4 for  $Y = J_{\varepsilon,y}$  and for  $x_- = (k_{y,+}^\varepsilon + 1)\ell$  and  $x_+ = (k_{z,-}^\varepsilon - 1)\ell$ . Because of the boundary conditions, the absence of layers to the left of  $x_-$  and the right of  $x_+$  implies in particular that  $u \leq 1 - \delta$  to the left of  $x_-$  and that  $u \geq -1 + \delta$  to the right of  $x_+$ . Hence we deduce that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_{y,3}) \lesssim L_\varepsilon \exp\left(-\frac{c_0 - \gamma}{\varepsilon}\right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{y,\varepsilon}). \quad (5.7)$$

The second lemma establishes that, on the other hand, when there are no extra layers, there is only a small probability of making an excursion from  $-1$  at some gridpoint in  $J_{y,-}^\varepsilon$  (respectively, an excursion from  $1$  at some gridpoint in  $J_{z,+}^\varepsilon$ ). The result from the second lemma is exactly the necessary ingredient that we need in Step 2 in order to invoke Lemma 5.2.

**Lemma 5.5.** *Fix any  $M < \infty$  sufficiently large. For any  $\delta > 0$  sufficiently small and  $\ell < \infty$  sufficiently large, there exists  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  we have*

$$\begin{aligned}\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathcal{J}_{y,\varepsilon} \cap \mathbb{C}\mathcal{A}_{y,3}^- : |u(x_k) + 1| \geq \frac{1}{2} \text{ for some } k \in \mathcal{I}_-^\varepsilon) \\ \leq \exp\left(-\frac{c_1}{2\varepsilon}\right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{y,\varepsilon}),\end{aligned} \quad (5.8)$$

and, similarly,

$$\begin{aligned}\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathcal{J}_{y,\varepsilon} \cap \mathbb{C}\mathcal{A}_{y,\varepsilon}^+ : |u(x_k) - 1| \geq \frac{1}{2} \text{ for some } k \in \mathcal{I}_+^\varepsilon) \\ \leq \exp\left(-\frac{c_1}{2\varepsilon}\right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{y,\varepsilon}),\end{aligned} \quad (5.9)$$

where  $c_1$  is defined in (6.86), below.

**Step 2.** The second step of the proof consists of showing that paths in the set  $\mathcal{J}_{y,\varepsilon}$  have hitting points of  $-1$  in  $J_{y,-}^\varepsilon$  and hitting points of  $+1$  in  $J_{z,+}^\varepsilon$  with large probability. This is captured by the following lemma, which is also proved in Subsection 6.6.

**Lemma 5.6.** *There exists  $C < \infty$  with the following property. Fix any  $\gamma > 0$  and any  $M < \infty$  sufficiently large. For any  $\delta > 0$  sufficiently small and  $\ell < \infty$*

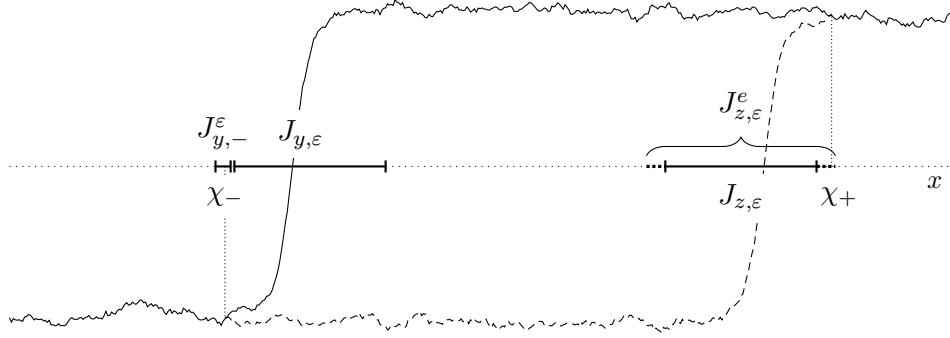


FIGURE 5.2. The reflection operator  $R_{y,z}$  performs a point reflection of the path between the left and right hitting points  $\chi_\pm$ . In this way the  $\delta^-$  transition in  $J_{y,\varepsilon}$  is mapped into  $J_{z,\varepsilon}^e$ .

sufficiently large, there exists  $\varepsilon_0 > 0$  and  $\lambda > 0$  such that, for all  $\varepsilon \leq \varepsilon_0$ , we have

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in \mathcal{J}_{y,\varepsilon} : \text{no hitting of } -1 \text{ in } J_{y,-}^\varepsilon \right) \\ \leq \frac{1}{2} E(\varepsilon) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}), \end{aligned} \quad (5.10)$$

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in \mathcal{J}_{y,\varepsilon} : \text{no hitting of } +1 \text{ in } J_{z,+}^\varepsilon \right) \\ \leq \frac{1}{2} E(\varepsilon) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}), \end{aligned} \quad (5.11)$$

where the error term satisfies

$$E(\varepsilon) := C \left( \lambda^{\bar{K}_\varepsilon} + L_\varepsilon \exp \left( -\frac{c_0 - \gamma}{\varepsilon} \right) + \exp \left( -\frac{c_1}{2\varepsilon} \right) \right), \quad (5.12)$$

and  $c_1$  is defined in (6.86), below.

**Step 3.** Now we are ready to define the reflection operator  $R_{y,z}$ . First, we define the following left and right stopping points

$$\begin{aligned} \chi_- &:= \inf \{ x \in J_{y,-}^\varepsilon : u(x) = -1 \}, \\ \chi_+ &:= \sup \{ x \in J_{z,+}^\varepsilon : u(x) = 1 \}. \end{aligned}$$

Here we use the convention that  $\chi_- = L_\varepsilon$  if there is no hitting point of  $-1$  in  $J_{y,-}^\varepsilon$  and similarly  $\chi_+ = -L_\varepsilon$  if there is no hitting point of  $1$  in  $J_{z,+}^\varepsilon$ . We use these hitting points to define the reflection operator

$$R_{y,z}u(x) := \begin{cases} u(x) & \text{for } x \leq \chi_-, \\ -u(\chi_- + \chi_+ - x) & \text{for } \chi_- < x < \chi_+, \\ u(x) & \text{for } x \geq \chi_+, \end{cases} \quad (5.13)$$

if  $\chi_- \leq \chi_+$ . We set  $R_{y,z}$  to be the identity otherwise. In other words the operator  $R_{y,z}$  performs a point reflection of the graphs of  $u$  between the left and right stopping points  $\chi_\pm$ . As in Step 4 of the proof of the upper bound in Theorem 1.5, one argues that the strong Markov property (3.10) implies that  $R_{y,z}$  leaves the measure  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  invariant. The action of the reflection operator is illustrated in Figure 5.2.

Assume that  $u \in \mathcal{J}_{y,\varepsilon}$  is a path that admits a hitting point of  $-1$  in  $J_{y,-}^\varepsilon$  and a hitting point of  $+1$  in  $J_{z,+}^\varepsilon$ . Recall that if  $u \in \mathcal{J}_{y,\varepsilon}$ , then  $u$  has a  $\delta^-$  up transition layer of length  $\leq 2\ell$  in  $J_{y,\varepsilon}$ . Under  $R_{y,z}$  the  $\delta^-$  up transition layer is mapped from  $J_{y,\varepsilon}$  to near  $J_{z,\varepsilon}$  and we would like to conclude that the reflected path is contained within  $\mathcal{J}_{z,\varepsilon}$ .

Unfortunately, the layer does not necessarily fall within  $J_{z,\varepsilon}$ . What is true is that there is a  $\delta^-$  up layer of length less than  $2\ell$  in the extended interval

$$J_{z,\varepsilon}^e := [z - d_\varepsilon - 3\ell - h_\varepsilon, z + d_\varepsilon + 3\ell + h_\varepsilon]. \quad (5.14)$$

(Recall that  $h_\varepsilon$ , the length of the auxiliary intervals, was defined above in (5.3).)

Let us denote by  $\mathcal{J}_{z,\varepsilon}^e$  the set of functions with a  $\delta^-$  up transition layer of length less than  $2\ell$  in  $J_{z,\varepsilon}^e$ :

$$\mathcal{J}_{y,\varepsilon}^e := \{u : u \text{ has a } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}^e\}.$$

In Step 2, we had established that

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in \mathcal{J}_{y,\varepsilon} : \text{no hitting of } -1 \text{ in } J_{y,-}^\varepsilon \text{ or no hitting of } +1 \text{ in } J_{z,+}^\varepsilon \right) \\ \leq E(\varepsilon) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}). \end{aligned}$$

Hence, as  $R_{y,z}$  leaves  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  invariant, we can conclude that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{z,\varepsilon}^e) \geq (1 - E(\varepsilon)) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}). \quad (5.15)$$

An analogous construction to turn transitions in  $J_{z,\varepsilon}$  into transitions near  $J_{y,\varepsilon}$  can be performed to obtain the same bound with  $J_{y,\varepsilon}$  and  $J_{z,\varepsilon}$  interchanged.

Step 4. In this step, we establish the bound (5.1). For notational convenience we will establish the bound in the case of the center interval  $[-d_\varepsilon, d_\varepsilon]$ , but our argument does not depend on this. More precisely, what we show is that for some  $\delta > 0$  and any  $\alpha > 0$ , there exists an  $\varepsilon_0 > 0$  such that, for  $\varepsilon \leq \varepsilon_0$ , we have

$$\begin{aligned} \left| \frac{L_\varepsilon}{d_\varepsilon} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (u \in \mathbb{C}A_1 : \text{at least one } \delta^- \text{ up layer} \right. \\ \left. \text{of length } \leq 2\ell \text{ in } [-d_\varepsilon, d_\varepsilon]) - 1 \right| \leq \alpha. \end{aligned} \quad (5.16)$$

The main ingredient will be the estimate (5.15). We will also use make use of Lemma 5.4 but except for that, the argument is completely elementary and only consists of choosing the right intervals and sets of paths.

We first split up the system into smaller blocks. Actually, it will useful to define two different partitions  $\{J_{k,\varepsilon}, k = -M_\varepsilon, \dots, M_\varepsilon - 1\}$  and  $\{J_{m,\varepsilon}^e, m = -\tilde{M}_\varepsilon, \dots, \tilde{M}_\varepsilon - 1\}$  of  $[-L_\varepsilon, L_\varepsilon]$ . The lengths of the intervals  $J_{k,\varepsilon}$  will be chosen small relative to  $d_\varepsilon$  but still large relative to  $|\log \varepsilon|$ . These intervals will be overlapping and play the role of  $J_{y,\varepsilon}$  when we apply (5.15). The intervals  $J_{m,\varepsilon}^e$  will be slightly larger than the intervals  $J_{k,\varepsilon}$  and will be of distance  $2\ell$  away from each other. They will be used as  $J_{z,\varepsilon}^e$  when applying (5.15).

We fix integers  $M_\varepsilon$  and  $k_\varepsilon$  such that

$$\begin{aligned} |\log \varepsilon| &\ll M_\varepsilon^{-1} L_\varepsilon \ll d_\varepsilon, \\ \text{and } k_\varepsilon M_\varepsilon^{-1} L_\varepsilon &\leq d_\varepsilon < (k_\varepsilon + 1) M_\varepsilon^{-1} L_\varepsilon. \end{aligned} \quad (5.17)$$



Then we set  $\tilde{d}_\varepsilon := L_\varepsilon/M_\varepsilon$  and define the overlapping intervals

$$J_{k,\varepsilon} := [k\tilde{d}_\varepsilon - 2\ell, (k+1)\tilde{d}_\varepsilon + 2\ell], \quad k = -(M_\varepsilon - 1), \dots, M_\varepsilon - 2.$$

The boundary intervals are defined as

$$J_{-M_\varepsilon,\varepsilon} := [-L_\varepsilon, -(M_\varepsilon - 1)\tilde{d}_\varepsilon + 2\ell] \quad \text{and} \\ J_{M_\varepsilon-1,\varepsilon} := [(M_\varepsilon - 1)\tilde{d}_\varepsilon - 2\ell, L_\varepsilon].$$

As above in (5.2), we then define the associated sets of paths as

$$\mathcal{J}_{k,\varepsilon} := \{u \in \mathbb{C}\mathcal{A}_1 : u \text{ has } \delta^- \text{ up layer of length } \leq 2\ell \text{ in } J_{k,\varepsilon}\}. \quad (5.18)$$

In order to define the slightly longer intervals, in analogy to the parameters  $h_\varepsilon$  and  $\bar{K}$  from Steps 1–3, we choose parameters  $\tilde{h}_\varepsilon$  and  $\tilde{K}_\varepsilon$  such that

$$\tilde{K}_\varepsilon \gg 1 \quad \text{and} \quad \tilde{h}_\varepsilon := \ell(2\tilde{K}_\varepsilon + 1)\tilde{K}_\varepsilon \ll \min \left\{ \tilde{d}, \exp \left( \frac{c_1}{4\varepsilon} \right) \right\}. \quad (5.19)$$

These parameters then define the error term  $E(\varepsilon)$  (see (5.12), above). Then we define the integers  $\tilde{M}_\varepsilon$  and  $m_\varepsilon$  such that

$$M_\varepsilon \left( \frac{1}{1 + (\tilde{h}_\varepsilon M_\varepsilon)L^{-1}} \right) - 1 \leq \tilde{M}_\varepsilon < M_\varepsilon \left( \frac{1}{1 + (\tilde{h}_\varepsilon M_\varepsilon)L^{-1}} \right) \\ \text{and} \quad m_\varepsilon \tilde{M}_\varepsilon^{-1} L_\varepsilon \leq d_\varepsilon < (m_\varepsilon + 1) \tilde{M}_\varepsilon^{-1} L_\varepsilon. \quad (5.20)$$

As above, we define the intervals

$$J_{m,\varepsilon}^e := [m\tilde{M}_\varepsilon^{-1} L_\varepsilon + 2\ell, (m+1)\tilde{M}_\varepsilon^{-1} L_\varepsilon - 2\ell], \quad m = -\tilde{M}_\varepsilon, \dots, \tilde{M}_\varepsilon - 1.$$

Each of these intervals  $J_{m,\varepsilon}^e$  is of length

$$\frac{L_\varepsilon}{\tilde{M}_\varepsilon} - 4\ell = \frac{L_\varepsilon}{M_\varepsilon} + \tilde{h}_\varepsilon - 4\ell, \quad (5.21)$$

and in particular these intervals are long enough to use them as  $J_{z,\varepsilon}^e$  in (5.15). Actually, when comparing (5.21) to (5.14), one notices a discrepancy in the length of  $10\ell$  but this can easily be treated by making  $\tilde{h}_\varepsilon$  a bit larger.

We define the associated sets of paths

$$\mathcal{J}_{m,\varepsilon}^e := \{u : u \text{ has a } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{m,\varepsilon}^e\}, \\ \mathcal{J}_{m,\varepsilon}^{e,*} := \{u \in \mathcal{J}_{m,\varepsilon}^e : u \text{ has no } \delta^- \text{ up layer in any } J_{n,\varepsilon}^e \text{ for any } n \neq m\}.$$

After these preliminary definitions, we are now ready to proceed to the proof of (5.16).

As mentioned above, the intervals  $J_{k,\varepsilon}$  are overlapping. In particular, every  $\delta^-$  layer  $\leq 2\ell$  in  $[-d_\varepsilon, d_\varepsilon]$  must be contained in at least one of the  $J_{k,\varepsilon}$ . This implies that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (u \in \mathbb{C}\mathcal{A}_1 : \delta^- \text{ up layer } \leq 2\ell \text{ in } [-d_\varepsilon, d_\varepsilon]) \\ \leq \sum_{k=-(k_\varepsilon+1)}^{k_\varepsilon} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{k,\varepsilon}). \quad (5.22)$$

In the same way we see that *every* possible path on  $[-L_\varepsilon, L_\varepsilon]$  must be either

- in one of the unlikely sets  $\mathcal{A}_1$  or  $\mathcal{A}_2$  defined above in (4.6) and (4.7)
- or in at least one of the sets  $\mathcal{J}_{k,\varepsilon}$ .

This implies that

$$\sum_{k=-M_\varepsilon}^{M_\varepsilon-1} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \geq 1 - \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1 \cup \mathcal{A}_2), \quad (5.23)$$

and hence we have

$$\max_k \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \geq \frac{1 - \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1 \cup \mathcal{A}_2)}{2M_\varepsilon}. \quad (5.24)$$

On the other hand, applying (5.15) gives that for any  $k$  and  $m$

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \leq (1 - E(\varepsilon))^{-1} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^e). \quad (5.25)$$

Then, applying Lemma 5.4 with  $Y = \mathcal{J}_{m,\varepsilon}^e$ , we have for every  $m$  that

$$(1 - E(\varepsilon)) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^e) \leq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}). \quad (5.26)$$

Finally, the sets  $\mathcal{J}_{m,\varepsilon}^{e,*}$  are all disjoint and in particular, we have

$$\sum_{m=-\tilde{M}_\varepsilon}^{\tilde{M}_\varepsilon-1} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}) \leq 1,$$

which implies that

$$\min_m \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}) \leq \frac{1}{2\tilde{M}_\varepsilon}. \quad (5.27)$$

We now collect ingredients to deduce the upper bound

$$\begin{aligned} & \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 : u \text{ has a } \delta^- \text{ up layer } \leq 2\ell \text{ in } [-d_\varepsilon, d_\varepsilon]) \\ & \stackrel{(5.22)}{\leq} \sum_{k=-(k_\varepsilon+1)}^{k_\varepsilon} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \\ & \leq (2k_\varepsilon + 2) \max_k \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \\ & \stackrel{(5.25)}{\leq} (1 - E(\varepsilon))^{-1} (2k_\varepsilon + 2) \min_m \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^e) \\ & \stackrel{(5.26)}{\leq} (1 - E(\varepsilon))^{-2} (2k_\varepsilon + 2) \min_m \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}) \\ & \stackrel{(5.27)}{\leq} (1 - E(\varepsilon))^{-2} \frac{(k_\varepsilon + 1)}{\tilde{M}_\varepsilon}. \end{aligned} \quad (5.28)$$

The proof of the lower bound now follows along similar lines:

$$\begin{aligned}
& \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 : u \text{ has a } \delta^- \text{ up layer } \leq 2\ell \text{ in } [-d_\varepsilon, d_\varepsilon]) \\
& \quad + \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1) \\
(5.20) \quad & \geq \sum_{m=-m_\varepsilon}^{m_\varepsilon-1} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}) \\
& \geq 2m_\varepsilon \min_m \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^{e,*}) \\
(5.26) \quad & \geq (1 - E(\varepsilon)) 2m_\varepsilon \min_m \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{m,\varepsilon}^e) \\
(5.25) \quad & \geq (1 - E(\varepsilon))^2 2m_\varepsilon \max_k \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{k,\varepsilon}) \\
(5.24) \quad & \geq (1 - E(\varepsilon))^2 (1 - \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1 \cup \mathcal{A}_2)) \frac{m_\varepsilon}{M_\varepsilon}. \tag{5.29}
\end{aligned}$$

Now, from the assumptions (5.17) on  $k_\varepsilon$  and  $M_\varepsilon$  as well as the assumptions (5.20) on  $m_\varepsilon$  and  $\tilde{M}_\varepsilon$ , we have that

$$1 \lesssim \frac{m_\varepsilon}{M_\varepsilon} \frac{L_\varepsilon}{d_\varepsilon} \leq 1 \quad \text{and} \quad 1 \leq \frac{k_\varepsilon + 1}{\tilde{M}_\varepsilon} \frac{L_\varepsilon}{d_\varepsilon} \lesssim 1. \tag{5.30}$$

Moreover, if we choose for instance  $M \geq 4C_2 c_0$  in the bound (4.9) on  $\mathcal{A}_1$ , we recover

$$L_\varepsilon \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1) \ll 1 \ll d_\varepsilon. \tag{5.31}$$

Combining (5.28), (5.29), (4.13), (5.30), and (5.31) establishes (5.16), as desired.

**Step 5.** It remains to remove the restriction on the length of the layer and improve from a  $\delta^-$  up layer to a full up layer.

The upper bound is immediate, since

$$\begin{aligned}
& \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 \text{ and there exists an up layer in } [y - d_\varepsilon, y + d_\varepsilon]) \\
& \leq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 \text{ and } \delta^- \text{ up layer in } [y - d_\varepsilon, y + d_\varepsilon]) \\
& \leq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 \text{ and } \delta^- \text{ up layer } \leq 2\ell \text{ in } \\
& \quad [y - d_\varepsilon, y + d_\varepsilon]) \\
& \quad + \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{C}\mathcal{A}_1 \text{ and } u \in [-1 + \delta, 1 - \delta] \text{ on } \\
& \quad \text{all of } [x_k, x_{k+1}] \text{ for some } k) \\
(5.1),(4.13) \quad & \leq (1 + \alpha) \frac{d_\varepsilon}{L_\varepsilon} + L_\varepsilon \exp\left(-\frac{\delta^2 \ell}{C_1 \varepsilon}\right) \\
& \leq (1 + 2\alpha) \frac{d_\varepsilon}{L_\varepsilon},
\end{aligned}$$

for  $\ell$  large enough with respect to  $1/\delta^2$ .

For the lower bound, on the other hand, we use Step 2 once more. To this end, we will consider layers falling strictly interior to  $J_{y,\varepsilon}$  on the subset  $J_{y,\varepsilon}^\circ := [y - d_\varepsilon + h_\varepsilon + 3\ell, y + d_\varepsilon - h_\varepsilon - 3\ell]$ . Then, according to Step 2, there is a high

probability of hitting  $\pm 1$  on  $J_{y,\varepsilon} \setminus J_{y,\varepsilon}^\odot$ . More precisely, notice that we can estimate

$$\begin{aligned}
& \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and there exists an up layer in } [y - d_\varepsilon, y + d_\varepsilon]) \\
& \geq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and there exists a } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}^\odot \\
& \quad \text{and } u \text{ hits } -1 \text{ in } (y - d_\varepsilon, y - d_\varepsilon + h_\varepsilon + 3\ell) \\
& \quad \text{and } u \text{ hits } +1 \text{ in } (y + d_\varepsilon - h_\varepsilon - 3\ell, y + d_\varepsilon)) \\
& \geq (1 - E(\varepsilon)) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and} \\
& \quad \text{there exists a } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}^\odot), \tag{5.32}
\end{aligned}$$

where in the last line, we have applied Lemma 5.6. On the other hand the probability on the last line can be estimated

$$\begin{aligned}
& \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}^\odot) \\
& \geq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}) \\
& \quad - \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and } \delta^- \text{ up layer } \leq 2\ell \\
& \quad \quad \text{in } (y - d_\varepsilon, y - d_\varepsilon + h_\varepsilon + 5\ell)) \\
& \quad - \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(u \in \mathbb{CA}_1 \text{ and } \delta^- \text{ up layer } \leq 2\ell \\
& \quad \quad \text{in } (y + d_\varepsilon - h_\varepsilon - 5\ell, y + d_\varepsilon)). \tag{5.33}
\end{aligned}$$

Applying the bound (5.1) to each term in (5.33) and substituting into (5.32) completes the lower bound.

Finally, recalling the bound (4.9) on the probability of  $\mathcal{A}_1$  completes the proof of Theorem 1.9.  $\square$

## 6. PROOFS OF THE LEMMAS

**6.1. Proofs of preliminary energy lemmas.** The energy lemmas rely on *upper bounds* and *lower bounds* for the energy over various sets. The upper bounds are derived based on constructions. (The minimum value of the energy is necessarily less than or equal to the value that we can achieve with any given construction.) The lower bound, on the other hand, describes the best possible value for *any function* and is based on the so-called Modica-Mortola trick discussed in Section 2. Before we begin, we make a remark about our constructions.

*Remark 6.1.* In addition to giving us an ODE for the energy minimizer on  $\mathbb{R}$ , equation (2.1) serves as the backbone for the *constructions* that are used to establish upper bounds for energy minimization problems on finite systems. For instance, suppose we want to minimize the energy on  $(-\ell, \ell)$  subject to  $u(\pm\ell) = \pm 1$ . For  $\ell$  large, we can build a construction that almost achieves the cost  $c_0$ . Specifically, consider the centered solution of (2.2) on  $(-\ell + a, \ell - a)$  for  $a = 1/\ell$ . Linearly interpolate from its value at  $-\ell + a$  to  $-1$  at  $-\ell$ , and symmetrically at the other end. Because of the exponential convergence of the minimizer to  $\pm 1$  (cf., Lemma 2.1), the energy on  $(-\ell, -\ell + a)$  and  $(\ell - a, \ell)$  is  $o(1)$  as  $\ell \uparrow \infty$ . Similarly, if we minimize the energy over functions satisfying  $u(\pm\ell) = \pm M$  for  $M$  large, we can build a piecewise-defined construction that goes from  $-M$  at  $-\ell$  to a neighborhood of  $-1$  at  $-\ell/2$ , goes from a neighborhood of  $-1$  at  $-\ell/2 + a$  to a neighborhood of 1

at  $\ell/2 - a$ , and goes from a neighborhood of 1 at  $\ell/2$  to  $M$  at  $\ell$ , with linear interpolation near  $\pm\ell/2$  to make the function continuous. The cost of such a construction is

$$\int_{-M}^{-1} \sqrt{2V(u)} du + \int_{-1}^1 \sqrt{2V(u)} du + \int_1^M \sqrt{2V(u)} du + o(1)_{\ell \uparrow \infty},$$

where we write the integrals separately to emphasize the additivity of the energy over the three subintervals described above. Because according to (2.1) we can get a good bound using increasing or decreasing functions, the analogous bounds hold for  $u(\pm\ell) = \mp M$ ,  $u(\pm\ell) = M$ , et cetera.

If  $M$  is very large, the constant  $\ell_*$  in the energy lemmas may also need to be very large in order to make the  $o(1)$  term small. The idea in all of the following proofs is to make this term small enough so that it can be absorbed into a  $\delta$ -dependent term, so the ordering of the constants is important: We fix  $M$  (large) and  $\delta$  (small) and then choose  $\ell_*$  large enough so that the term(s) that are  $o(1)$  with respect to  $\ell$  can be absorbed.

In what follows, it will be convenient to introduce the notation:

$$\varphi_{-1}(u) = \left| \int_{-1}^u \sqrt{2V(s)} ds \right|, \quad \varphi_{+1}(u) = \left| \int_u^1 \sqrt{2V(s)} dx \right|.$$

*Proof of Lemma 2.3.* We will establish (2.4) via an upper bound on the energy over  $\mathcal{A}^{\text{bc}}$  and a lower bound on the energy over  $\mathcal{A}_0^{\text{bc}}$ . Because of the extra condition in  $\mathcal{A}_0^{\text{bc}}$ , the energy on  $(-\ell, \ell)$  is large (of order  $\delta^2 \ell$ ), and we do not have to be as careful about the boundary conditions as usual. A rough bound will suffice.

Step 1. As explained in Remark 6.1, the upper bound relies on a construction. Given any  $u_- \in [-M, M]$ , we can use the solution of (2.1) to connect to a neighborhood of 1 or  $-1$ , and similarly for  $u_+$ . If the optimal connection for  $u_-$  is to  $-1$  and the optimal connection for  $u_+$  is to  $+1$ , then in order to build a continuous construction, we incur the additional cost  $\varphi_{-1}(1) = c_0$ , where we have used the notation introduced above and recalled the value of  $c_0$  from (1.9). (If the optimal connection for  $u_-$  and  $u_+$  is to the same value, then the construction does not incur this extra cost, but the upper bound is still valid.) Putting together these three pieces of the construction and the small correction terms for continuity (see Remark 6.1), we can express the upper bound derived in this way as:

$$\begin{aligned} \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) &\leq \min\{\varphi_{-1}(u_-), \varphi_{+1}(u_-)\} \\ &\quad + \min\{\varphi_{-1}(u_+), \varphi_{+1}(u_+)\} + c_0 + o(1)_{\ell \uparrow \infty}. \end{aligned} \quad (6.1)$$

Note that Assumption 1.1 allows that  $o(1)_{\ell \uparrow \infty}$  may depend on  $M$ : If  $u_-$  is very large, the (near) optimal connection from  $u_-$  to  $1$  requires a lot of space. This explains why  $\ell_*$  in the statement of the lemma depends on  $M$ .

Step 2. Now we turn to the lower bound over  $\mathcal{A}_0^{\text{bc}}$ . On the one hand, on  $(-\ell, \ell)$ , the condition in  $\mathcal{A}_0^{\text{bc}}$  implies that the integral of  $V$  over  $(-\ell, \ell)$  cannot be too small. Using the quadratic behavior of  $V$  near  $\pm 1$  (see Assumption 1.1), we have for  $\delta$  small enough

$$E_{(-\ell, \ell)}(u) \geq \int_{-\ell}^{\ell} V(u) du \geq \frac{V''(1) \ell \delta^2}{2}. \quad (6.2)$$

To integrate over the rest of the interval, we recall the trick of Modica and Mortola that was explained in Section 2. Consider first  $(-2\ell, -\ell)$ . We divide into two cases:  $|u_-| > 1$  and the complement.

If  $u \in \mathcal{A}_0^{\text{bc}}$  and  $|u_-| > 1$ , then there is a point  $x_- \in (-2\ell, -\ell)$  such that  $|u(x_-)| = 1$ . In this case, the Modica-Mortola trick on  $(-2\ell, x_-)$  gives

$$E_{(-2\ell, -\ell)}(u) \geq E_{(-2\ell, x_-)}(u) \geq \min\{\varphi_{-1}(u_-), \varphi_{+1}(u_-)\}. \quad (6.3)$$

On the other hand if  $|u_-| \leq 1$ , then for  $\ell$  large enough, we have

$$\min\{\varphi_{-1}(u_-), \varphi_{+1}(u_-)\} \leq \frac{V''(1) \ell \delta^2}{8}. \quad (6.4)$$

If  $|u_-| > 1$ , then adding the contributions from (6.2) and (6.3) and subtracting the contribution from (6.1) gives

$$\begin{aligned} & \inf_{u \in \mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \\ & \geq \frac{V''(1) \ell \delta^2}{2} - \min\{\varphi_{-1}(u_+), \varphi_{+1}(u_+)\} - c_0 + o(1)_{\ell \uparrow \infty}. \end{aligned}$$

On the other hand if  $|u_-| \leq 1$ , then the contributions from (6.2) and (6.1) together with the bound from (6.4) imply

$$\begin{aligned} & \inf_{u \in \mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \\ & \geq \frac{3V''(1) \ell \delta^2}{8} - \min\{\varphi_{-1}(u_+), \varphi_{+1}(u_+)\} - c_0 + o(1)_{\ell \uparrow \infty}. \end{aligned}$$

Since this is a weaker bound, it holds in either case.

Repeating the identical argument on  $(\ell, 2\ell)$  and in addition absorbing  $c_0$  by  $V''(1)\ell\delta^2/8$  gives

$$\begin{aligned} & \inf_{u \in \mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \\ & \geq \frac{V''(1) \ell \delta^2}{8} + o(1)_{\ell \uparrow \infty}, \end{aligned}$$

which completes the proof of Lemma 2.3.  $\square$

*Proof of Lemma 2.5.* We rewrite the set  $\mathcal{A}_0^{\text{bc}}$  as

$$\mathcal{A}_0^{\text{bc}} = \mathcal{A}_- \cup \mathcal{A}_+,$$

where the  $\mathcal{A}_{\pm}$  are the sets of paths that perform a wasted excursion starting from a neighborhood of  $\pm 1$ . We will prove the bound on the energy difference for  $\mathcal{A}_+$ . The corresponding bound for  $\mathcal{A}_-$  follows in the same way.

As usual, our task is to produce appropriate upper and lower bounds.

Step 1. The upper bound on  $\inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u)$  is by construction. Consider the function  $\bar{u}$  that minimizes  $E_{(-2\ell, 2\ell)}$  subject to

$$u(\pm 2\ell) = u_{\pm}, \quad u(0) = 1,$$

and notice that

$$\begin{aligned} E_{(-2\ell, 2\ell)}(\bar{u}) &= \inf \left\{ \int_{-2\ell}^0 \frac{1}{2} (\partial_x u)^2 + V(u) dx : u(0) = 1, u(-2\ell) = u_- \right\} \\ &\quad + \inf \left\{ \int_0^{2\ell} \frac{1}{2} (\partial_x u)^2 + V(u) dx : u(0) = 1, u(2\ell) = u_+ \right\} \\ &= \varphi_{+1}(u_-) + \varphi_{+1}(u_+) + o(1)_{\ell \uparrow \infty}, \end{aligned}$$

uniformly for  $u_{\pm} \in [-M, M]$ . (This can be established by building a construction by hand, as we have explained in Remark 6.1 and the proof of Lemma 2.3.) Hence, since  $\bar{u} \in \mathcal{A}^{\text{bc}}$ , we have the (not necessarily tight) upper bound

$$\inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \leq E_{(-2\ell, 2\ell)}(\bar{u}) = \varphi_{+1}(u_-) + \varphi_{+1}(u_+) + o(1)_{\ell \uparrow \infty}. \quad (6.5)$$

Step 2. We now turn to the lower bound on  $\inf_{\mathcal{A}_+} E_{(-2\ell, 2\ell)}(u)$ . Recall the points  $x_{\pm}$  that follow from the definition of  $\mathcal{A}_+$  and Definition 2.4. Because of the properties of the potential, we may without loss of generality assume that  $u(x_{\pm}) = 1 - \delta$  and  $u(x_0) = \delta$ .

We now use the Modica-Mortola trick on  $(-2\ell, x_-) \cup (x_+, 2\ell)$  to recover

$$\begin{aligned} &E_{(-2\ell, x_-)}(u) + E_{(x_+, 2\ell)}(u) \\ &\geq \varphi_{+1}(u_-) + \varphi_{+1}(u_+) - C\delta^2 \\ &\stackrel{(6.5)}{\geq} \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - C\delta^2 - o(1)_{\ell \uparrow \infty}. \end{aligned} \quad (6.6)$$

On the other hand, applying the Modica-Mortola trick on  $(x_-, x_0) \cup (x_0, x_+)$  gives

$$E_{(x_-, x_0)}(u) + E_{(x_0, x_+)}(u) \geq 2 \int_{\delta}^{1-\delta} \sqrt{2V(u)} du \stackrel{(1.9)}{=} c_0 - C\delta. \quad (6.7)$$

Combining (6.6) and (6.7) completes the proof of Lemma 2.5.  $\square$

*Proof of Lemma 2.8. Step 1.* For the upper bound over  $\mathcal{A}_{\delta, pre}^{\text{bc}}$ , we use the function  $\bar{u}$  that minimizes the energy subject to

$$u(\pm 2\ell) = u_{\pm}, \quad u(\pm \ell) = -1 - 2\delta, \quad u(0) = \delta.$$

As in the proof of Lemma 2.5, we observe that  $\bar{u} \in \mathcal{A}_{\delta, pre}^{\text{bc}}$  and hence the construction gives an upper bound

$$\begin{aligned} \inf_{\mathcal{A}_{\delta, pre}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) &\leq \varphi_{-1}(u_-) + \varphi_{-1}(u_+) + 2\varphi_{+1}(0) + C\delta + o(1)_{\ell \uparrow \infty} \\ &\stackrel{(1.9)}{=} \varphi_{-1}(u_-) + \varphi_{-1}(u_+) + c_0 + C\delta + o(1)_{\ell \uparrow \infty}. \end{aligned} \quad (6.8)$$

Step 2. For the lower bound over  $\mathcal{A}^{\text{bc}}$ , we observe that for any  $u \in \mathcal{A}^{\text{bc}}$ , either there is a point  $x_- \in (-2\ell, 0)$  and a point  $x_+ \in (0, 2\ell)$  such that  $u(x_{\pm})$  is in a  $\delta$  neighborhood of 1 or  $-1$ , or else the energy (by the same argument as in the proof of Lemma 2.3) is bounded below by  $\delta^2 \ell V''(1)/2$  for  $\delta$  small enough. We can choose  $\ell$  so large that this is greater than  $\varphi_{-1}(u_-) + \varphi_{-1}(u_+)$  and hence dominates the boundary terms in (6.8). On the other hand, if the points  $x_{\pm}$  exist, then by the

usual trick of Modica and Mortola, we recover

$$\begin{aligned} \inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) &\geq \min\{\varphi_{-1}(u_-), \varphi_{+1}(u_-)\} + \min\{\varphi_{-1}(u_+), \varphi_{+1}(u_+)\} - C\delta \\ &= \varphi_{-1}(u_-) + \varphi_{-1}(u_+) - C\delta, \end{aligned} \quad (6.9)$$

where the second line follows by virtue of the boundary conditions  $u_{\pm} \in [-M, 0]$  and the symmetry of the potential.

The combination of (6.8) and (6.9) completes the proof of Lemma 2.8.  $\square$

## 6.2. Proof of the strong Markov property.

*Proof of Lemma 3.1.* By subtracting  $h_{(x_-, x_+)}^{u_-, u_+}$ , we can reduce the problem to the case of zero boundary conditions. Under  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$ ,  $u - u_{\hat{x}_-}^{\hat{x}_+}$  and  $u_{\hat{x}_-}^{\hat{x}_+}$  are jointly Gaussian and centered, because they are both linear images of  $u$ . So it is sufficient to calculate their covariances. Using (3.1), it is easy to see that, for all  $x_1, x_2 \in [x_-, x_+]$ , one has

$$\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, 0, 0}} \left( (u - u_{\hat{x}_-}^{\hat{x}_+})(x_1) u_{\hat{x}_-}^{\hat{x}_+}(x_2) \right) = 0,$$

and for  $x_1, x_2 \in [\hat{x}_-, \hat{x}_+]$ , one has

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, 0, 0}} \left( (u - u_{\hat{x}_-}^{\hat{x}_+})(x_1) (u - u_{\hat{x}_-}^{\hat{x}_+})(x_2) \right) \\ = \frac{\varepsilon}{\hat{x}_+ - \hat{x}_-} \left( (x_1 - \hat{x}_-)(\hat{x}_+ - x_2) \wedge (x_2 - \hat{x}_-)(\hat{x}_+ - x_1) \right). \end{aligned}$$

This shows the claim.  $\square$

*Proof of Lemma 3.2.* We start by observing that the statement of Lemma 3.1 implies that

$$\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}} (\Phi | \mathcal{F}_{[x_-, \hat{x}_-]} \vee \mathcal{F}_{[\hat{x}_+, x_+]}) = \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mathcal{W}_{\varepsilon, \mathbf{u}}} (\Phi). \quad (6.10)$$

In order to prove the desired statement (3.7), observe that the density of  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  with respect to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  can be written as

$$\exp \left( - \frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) = \Psi_- \Psi_{\odot} \Psi_+,$$

where

$$\Psi_- := \exp \left( - \frac{1}{\varepsilon} \int_{x_-}^{\hat{x}_-} V(u) dx \right), \quad \Psi_+ := \exp \left( - \frac{1}{\varepsilon} \int_{\hat{x}_+}^{x_+} V(u) dx \right),$$

$$\text{and } \Psi_{\odot} := \exp \left( - \frac{1}{\varepsilon} \int_{\hat{x}_-}^{\hat{x}_+} V(u) dx \right)$$

are measurable with respect to  $\mathcal{F}_{[x_-, \hat{x}_-]}$ ,  $\mathcal{F}_{[\hat{x}_-, \hat{x}_+]}$ , and  $\mathcal{F}_{[\hat{x}_+, x_+]}$ . Suppose that test functions  $\Xi_-$  and  $\Xi_+$  are measurable with respect to  $\mathcal{F}_{[x_-, \hat{x}_-]}$  and  $\mathcal{F}_{[\hat{x}_+, x_+]}$ . Then



we get

$$\begin{aligned}
& \mathbb{E}_{(x-, x_+)}^{\mu_\varepsilon, u_-, u_+} \left( \Phi \Xi_- \Xi_+ \right) \\
&= \frac{1}{\mathcal{Z}_{\varepsilon, (x-, x_+)}^{u_-, u_+}} \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Xi_- \Psi_- \Phi \Psi_\odot \Xi_+ \Psi_+ \right) \\
&\stackrel{(6.10)}{=} \frac{1}{\mathcal{Z}_{\varepsilon, (x-, x_+)}^{u_-, u_+}} \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Xi_- \Psi_- \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mathcal{W}_\varepsilon, \mathbf{u}} \left( \Phi \Psi_\odot \right) \Xi_+ \Psi_+ \right) \\
&= \frac{1}{\mathcal{Z}_{\varepsilon, (x-, x_+)}^{u_-, u_+}} \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Xi_- \Psi_- \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mathcal{W}_\varepsilon, \mathbf{u}} \left( \Psi_\odot \right) \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mu_\varepsilon, \mathbf{u}} \left( \Phi \right) \Xi_+ \Psi_+ \right) \\
&\stackrel{(6.10)}{=} \mathbb{E}_{(x-, x_+)}^{\mu_\varepsilon, u_-, u_+} \left( \Xi_- \mathbb{E}_{(\hat{x}_-, \hat{x}_+)}^{\mu_\varepsilon, \mathbf{u}} \left( \Phi \right) \Xi_+ \right).
\end{aligned} \tag{6.11}$$

This finishes the proof of Lemma 3.2.  $\square$

We are now ready to give a proof of the strong Markov property.

*Proof of Lemma 3.3:* We treat only the Gaussian case (3.9). Equation (3.10) then follows as in the proof of Lemma 3.2.

We start by proving (3.9) in the case in which  $\chi_-$  and  $\chi_+$  are left and right stopping points that attain values in a finite set  $\{\chi^1, \dots, \chi^N\}$ . Then we can write

$$\begin{aligned}
& \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Phi | \mathcal{F}_{[x-, \chi_-]} \vee \mathcal{F}_{[\chi_+, x_+]} \right) \\
&= \sum_{n=1}^N \sum_{m=1}^N \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Phi \mathbf{1}_{\{\chi_- = \chi^n\}} \mathbf{1}_{\{\chi_+ = \chi^m\}} | \mathcal{F}_{[x-, \chi_-]} \vee \mathcal{F}_{[\chi_+, x_+]} \right) \\
&= \sum_{n=1}^N \sum_{m=1}^N \mathbf{1}_{\{\chi_- = \chi^n\}} \mathbf{1}_{\{\chi_+ = \chi^m\}} \mathbb{E}_{(x-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \Phi | \mathcal{F}_{[x-, \chi^n]} \vee \mathcal{F}_{[\chi^m, x_+]} \right) \\
&\stackrel{(6.10)}{=} \sum_{n=1}^N \sum_{m=1}^N \mathbf{1}_{\{\chi_- = \chi^n\}} \mathbf{1}_{\{\chi_+ = \chi^m\}} \mathbb{E}_{(\chi^n, \chi^m)}^{\mathcal{W}_\varepsilon, \mathbf{u}} (\Phi) \\
&= \mathbb{E}_{(\chi_-, \chi_+)}^{\mathcal{W}_\varepsilon, \mathbf{u}} (\Phi).
\end{aligned}$$

In the second equality, we have used the fact that the  $\chi_\pm$  are left and right stopping points.

In order to see the general case, we approximate the stopping points by

$$\begin{aligned}
\chi_-^N &:= \inf \{ x = i 2^{-N} : i \in \mathbb{Z}, x \geq \chi_- \}, \\
\chi_+^N &:= \sup \{ x = i 2^{-N} : i \in \mathbb{Z}, x \leq \chi_+ \}.
\end{aligned}$$

Then  $\chi_-^N$  and  $\chi_+^N$  are stopping points taking values in a finite set and, in particular, (3.9) holds for them. We have

$$\chi_-^N \downarrow \chi_- \quad \text{and} \quad \chi_+^N \uparrow \chi_+ \quad \text{as } N \uparrow \infty.$$

Now, in order to conclude that (3.9) also holds for  $\chi_{\pm}$ , we first observe that for any continuous, bounded  $\Phi: C([x_-, x_+]) \rightarrow \mathbb{R}$ , we have for every path  $u$  that

$$\begin{aligned} & \mathbb{E}_{(\chi_-, \chi_+)}^{\mathcal{W}_{\varepsilon}, \mathbf{u}}(\Phi) \\ &= \lim_{N \rightarrow \infty} \mathbb{E}_{(\chi_-^N, \chi_+^N)}^{\mathcal{W}_{\varepsilon}, \mathbf{u}}(\Phi) \\ &= \lim_{N \rightarrow \infty} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon}, u_-, u_+}(\Phi | \mathcal{F}_{[x_-, \chi_-^N]} \vee \mathcal{F}_{[\chi_+^N, x_+]}) \\ &= \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon}, u_-, u_+}(\Phi | \mathcal{F}_{[x_-, \chi_-]} \vee \mathcal{F}_{[\chi_+, x_+]}). \end{aligned}$$

In the first step, we have used that, due to the continuity of  $u$ , the measures  $\mathcal{W}_{\varepsilon, (\chi_-^N, \chi_+^N)}^u$  converge weakly to  $\mathcal{W}_{\varepsilon, (\chi_-, \chi_+)}^u$ , as can easily be confirmed. In order to see the last line, it suffices to check that the limit in the third line does indeed satisfy the characteristic properties of a conditional expectation.

This equality can then be extended to arbitrary test functions  $\Phi$  with a standard monotone class argument (see e.g. [RY99, Ch. 0, Thm 2.2]).  $\square$

**6.3. Proof of large deviation bounds.** The large deviation bounds (3.17) and (3.18) are statements about the quotient of expectations of the form

$$\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon}, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right),$$

see (1.5). Consequently, the results will follow as soon as we establish upper and lower bounds on these expectations. Throughout this subsection,  $\mathcal{A}$  will always denote a set of continuous paths  $u$  on  $[x_-, x_+]$  that satisfy the boundary conditions  $u(x_{\pm}) = u_{\pm}$ , and topological notions like open or closed will always refer to the topology of uniform convergence. We will frequently use  $I_{x_-, x_+}(u)$ , the Gaussian energy of a path (defined in (3.2)), and  $I_{x_{\pm}}^{u_{\pm}}$ , the minimal Gaussian energy given the boundary conditions (defined in (3.15)).

The upper bound for the Gaussian expectation can then be stated as follows.

**Lemma 6.2** (Upper bound). *Fix constants  $M < \infty$ ,  $0 < \ell_- < \ell_+ < \infty$  and  $R < \infty$ . Suppose that  $\ell = (x_+ - x_-) \in [\ell_-, \ell_+]$  and  $u_{\pm} \in [-M, M]$ . Then for any  $\delta, \gamma > 0$ , there exists an  $\varepsilon_0 > 0$  such that for any measurable set  $\mathcal{A}$  satisfying*

$$\inf_{u \in B(\mathcal{A}, \delta)} E(u) - I_{x_{\pm}}^{u_{\pm}} \leq R \quad (6.12)$$

and for any  $\varepsilon \leq \varepsilon_0$ , we have

$$\begin{aligned} & \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon}, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(\mathcal{A}, \delta)} E(u) - I_{x_{\pm}}^{u_{\pm}} - \gamma \right) \right). \end{aligned} \quad (6.13)$$

Here  $\varepsilon_0$  depends on  $M, \ell_{\pm}, \sup_{|v| \leq M + \sqrt{2^{-1}(\ell_+ R + 1)} + 1} |V'(v)|, \delta$ , and  $\gamma$  but not on the particular choice of  $x_{\pm}, u_{\pm}$ , and it depends on  $\mathcal{A}$  only through condition (6.12).

As usual in large deviation theory, the derivation of lower bounds for integrals is reduced to the case of a ball

$$B(u_*, \delta) := \{u: \|u - u_*\|_{\infty} \leq \delta\}$$

around a suitably chosen profile  $u_*$ .

**Lemma 6.3** (Lower bound). *Fix constants  $M$  and  $\ell_+ < \infty$ . Suppose that  $\ell = x_+ - x_- \leq \ell_+$ ,  $u_{\pm} \in [-M, M]$ . Then for any profile  $u_*$  with*

$$\sup_{x \in (x_-, x_+)} |u_*(x)| \leq M \quad (6.14)$$

*and any  $\delta, \gamma > 0$ , there exists an  $\varepsilon_0 > 0$  such that for  $\varepsilon \leq \varepsilon_0$*

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}} \left( \mathbf{1}_{B(u_*, \delta)}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \exp \left( -\frac{1}{\varepsilon} (E(u_*) - I_{x_{\pm}}^{u_{\pm}} + \gamma) \right). \end{aligned} \quad (6.15)$$

*Here  $\varepsilon_0$  depends on  $\sup_{|v| \leq M+1} |V'(v)|, \ell_+, \delta$ , and  $\gamma$  but not on the particular choice of  $x_{\pm}, u_{\pm}$  and it depends on  $u_*$  only through the condition (6.14).*

Now we give the proofs of Lemmas 6.2 and 6.3. The proofs of Propositions 3.4 and 3.5 are given afterwards.

In order to prove the upper bound, we will invoke the known upper bound for Gaussian large deviations. In the current context, this can be stated as follows.

**Proposition 6.4** (Gaussian large deviation, see e.g. [Bog98, Cor. 4.9.3]). *For every closed set  $\mathcal{A}$  and for any  $\gamma > 0$ , there exists an  $\varepsilon_0 > 0$  such that for every  $\varepsilon \leq \varepsilon_0$  we have*

$$\mathcal{W}_{\varepsilon, (0,1)}^{0,0}(\mathcal{A}) \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in \mathcal{A}} I_{0,1}(u) - \gamma \right) \right). \quad (6.16)$$

The argument for Lemma 6.2 is an adaptation of the proof of [dH00, p. 34].

*Proof of Lemma 6.2. Step 1.* We start by reducing the general problem to the case of homogeneous boundary conditions on  $[0, 1]$ . To this end, we introduce the following affine transformation. We define the transformation  $T : u \mapsto \hat{u}$ , where for a given path  $u : [x_-, x_+] \rightarrow \mathbb{R}$  we denote by  $\hat{u} \in C([0, 1])$  the function

$$\hat{u}(x) := u(x_- + \ell x) - h_{0,1}^{u_-, u_+}(x). \quad (6.17)$$

Recall from (3.4) that  $h_{0,1}^{u_-, u_+}(x) = xu_+ + (1-x)u_-$ . It is clear that  $T$  is a bijection between the set of continuous paths  $u$  on  $[x_-, x_+]$  with boundary conditions  $u(x_{\pm}) = u_{\pm}$  and  $C([0, 1])$ , the space of continuous paths on  $[0, 1]$  with homogeneous boundary conditions. Furthermore, if  $u$  is distributed according to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ , then  $\hat{u}$  is distributed according to  $\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}$ . Note that the variance changes due to the rescaling by  $\ell$ .

The expectation that we want to bound can be expressed in terms of  $\hat{u}$  as

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ = \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}} \left( \mathbf{1}_{\hat{\mathcal{A}}}(\hat{u}) \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(\hat{u} + h_{0,1}^{u_-, u_+}) dx \right) \right), \end{aligned} \quad (6.18)$$

where  $\hat{\mathcal{A}} := \{Tu : u \in \mathcal{A}\}$ . On the other hand, the condition (6.12) and the right-hand side of the desired bound (6.13) can also be expressed in terms of  $\hat{u}$ , as we will now do. We have for every  $u$  that

$$E(u) = E_{\ell}^{u_{\pm}}(\hat{u}) + I_{x_{\pm}}^{u_{\pm}},$$

where, for convenience, we have introduced the notation

$$E_\ell^{u\pm}(\hat{u}) := \int_0^1 \frac{1}{2\ell} (\partial_x \hat{u})^2 + \ell V(\hat{u} + h_{0,1}^{u-,u+}) dx.$$

(Note that we have not included  $I_{x_\pm}^{u\pm}$  in the definition of the rescaled energy  $E_\ell^{u\pm}$ , because this way  $E_\ell^{u\pm}$  will appear as the natural rate functional.) Condition (6.12) can now be expressed as

$$\inf_{u \in B(\hat{\mathcal{A}}, \delta)} E_\ell^{u\pm}(\hat{u}) \leq R, \quad (6.19)$$

and for the right-hand side of (6.13), we get

$$\begin{aligned} & \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(\mathcal{A}, \delta)} E(u) - I_{x_\pm}^{u\pm} - \gamma \right) \right) \\ &= \exp \left( -\frac{1}{\varepsilon} \left( \inf_{\hat{u} \in B(\hat{\mathcal{A}}, \delta)} E_\ell^{u\pm}(\hat{u}) - \gamma \right) \right). \end{aligned}$$

Relabelling  $\hat{\mathcal{A}}$  as  $\mathcal{A}$  and  $\hat{u}$  as  $u$ , we conclude that it suffices to show that for every set  $\mathcal{A} \subseteq C([0, 1])$  satisfying

$$\inf_{u \in B(\mathcal{A}, \delta)} E_\ell^{u\pm}(u) \leq R, \quad (6.20)$$

we have for  $\varepsilon \leq \varepsilon_0$  that

$$\begin{aligned} & \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon,0,0}} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(u + h_{0,1}^{u\pm}) dx \right) \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(\mathcal{A}, \delta)} E_\ell^{u\pm}(u) - \gamma \right) \right). \end{aligned} \quad (6.21)$$

This bound will be established in Steps 2-4.

Step 2. The strategy to prove (6.21) consists of decomposing  $C([0, 1])$  into a set of paths with high Gaussian energy and a finite number of small balls with lower Gaussian energy. One can use the Gaussian large deviation bound (6.16) to bound the probability of the set of high Gaussian energy, which we will make to be a term of higher exponential order by choosing the Gaussian energy high enough. Then for the balls with lower Gaussian energy, the expectation over a given ball can be estimated by bounding an exponential factor by its supremum on that ball, and then bounding the Gaussian probability of the set using (6.16) again. Finally, one has to sum over all the balls. As the total number of balls is finite and the bounds decay exponentially, the largest of the summands determines the behavior.

The main difference with respect to the classical argument in [dH00] is that we choose a partition of  $C([0, 1])$  into sets that do not depend on  $\mathcal{A}$ . This is necessary to ensure that the number of balls is independent of  $\mathcal{A}$ . The price we have to pay is that on the right-hand side of (6.21) we take the infimum over the small neighborhood  $B(\mathcal{A}, \delta)$  of  $\mathcal{A}$  instead of taking it over  $\mathcal{A}$  only, as in the classical argument.

Let us now give the details: First, fix a  $\gamma < 1$  and let

$$\tilde{\delta} := \gamma \left( \ell_+ \sup_{|v| \leq \sqrt{2^{-1}(\ell_+ R + 1)} + M + 1} |V'(v)| \right)^{-1} \wedge \delta \wedge \frac{1}{2}. \quad (6.22)$$

The sublevel set

$$\mathcal{K}_{\ell_+ R} := \left\{ u : I_{0,1} \leq \ell_+ R \right\}$$

is compact in  $C([0, 1])$ , and we can cover it by a finite number  $N_{\tilde{\delta}, \ell_+ R}$  of open balls  $B(u_k, \tilde{\delta})$  of radius  $\tilde{\delta}$ , where  $u_k \in \mathcal{K}_{\ell_+ R}$  for each  $k$ . Note that  $\mathcal{A}$  does not enter here, so both the profiles  $u_k$  and the number  $N_{\tilde{\delta}, \ell_+ R}$  depend only on  $\gamma, \delta, \ell_+ R$ , and  $\sup_{|v| \leq \sqrt{2^{-1}(\ell_+ R + 1)} + M + 1} |V'(v)|$ , not on the set  $\mathcal{A}$  or the specific choice of  $x_{\pm}, u_{\pm}$ . Actually, it can be checked using the Hölder continuity of functions with bounded  $H^1$ -norm that this number grows like  $\exp(C(R\ell_+ \tilde{\delta}^{-1})^2)$ .

Using this covering and the positivity of  $V$ , we have for any set  $\mathcal{A}$  that

$$\begin{aligned} & \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon,0,0}} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(u + h_{0,1}^{u_{\pm}}) dx \right) \right) \\ & \leq \sum_{k=1}^{N_{\tilde{\delta}, \ell_+ R}} \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon,0,0}} \left( \mathbf{1}_{B(u_k, \tilde{\delta}) \cap \mathcal{A}}(u) \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(u + h_{0,1}^{u_{\pm}}) dx \right) \right) \\ & \quad + \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0} \left( \mathcal{A} \setminus \cup_k B(u_k, \tilde{\delta}) \right). \end{aligned} \quad (6.23)$$

Step 3. The last term in (6.23) can now easily be bounded:

$$\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0} \left( \mathcal{A} \setminus \cup_k B(u_k, \tilde{\delta}) \right) \leq \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0} \left( \mathbb{C} \cup_k B(u_k, \tilde{\delta}) \right). \quad (6.24)$$

The set  $\mathcal{B} := \mathbb{C} \cup_k B(u_k, \tilde{\delta})$  is closed and by definition  $\inf_{u \in \mathcal{B}} I_{0,1}(u) \geq \ell_+ R$ . Hence, the Gaussian large deviation bound (6.16) implies that there exists an  $\tilde{\varepsilon}_0 > 0$  such that, for  $\varepsilon \leq \tilde{\varepsilon}_0$ , we have

$$\mathcal{W}_{\varepsilon, (0,1)}^{0,0}(\mathcal{B}) \leq \exp \left( -\frac{1}{\varepsilon} (\ell_+ R - \gamma) \right).$$

Now we choose  $\varepsilon_0 = \tilde{\varepsilon}_0 \ell_+^{-1}$ . Then, for  $\varepsilon \leq \varepsilon_0$ , we can conclude that

$$\begin{aligned} & \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0} \left( \mathcal{A} \setminus \cup_k B(u_k, \tilde{\delta}) \right) \\ & \leq \exp \left( -\frac{1}{\ell\varepsilon} (\ell_+ R - \gamma) \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} \left( R - \frac{\gamma}{\ell_-} \right) \right) \\ & \stackrel{(6.19)}{\leq} \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(\mathcal{A}, \delta)} E_{\ell}^{u_{\pm}}(u) - \frac{\gamma}{\ell_-} \right) \right). \end{aligned} \quad (6.25)$$

Step 4. It remains to bound the sum on the right-hand side of (6.23). Since the number of summands  $N_{\tilde{\delta}, \ell_+ R}$  remains constant as  $\varepsilon \downarrow 0$ , the sum is dominated by the largest summand. Specifically, after fixing  $\gamma, \delta, \ell_+ R$  and  $M$ , we can choose  $\varepsilon_0 > 0$  sufficiently small so that  $\varepsilon \leq \varepsilon_0$  implies

$$N_{\tilde{\delta}, \ell_+ R} = \exp \left( \frac{1}{\varepsilon} \left( \varepsilon \log(N_{\tilde{\delta}, \ell_+ R}) \right) \right) \leq \exp \left( \frac{\gamma}{\varepsilon} \right). \quad (6.26)$$

Hence, up to an extra factor of  $\gamma$ , it is sufficient to obtain a good exponential bound on the largest summand on the right-hand side of (6.23).

If  $B(u_k, \tilde{\delta}) \cap \mathcal{A}$  is empty, the largest summand is zero. Otherwise, we have

$$\begin{aligned} & \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon,0,0}} \left( \mathbf{1}_{B(u_k, \tilde{\delta}) \cap \mathcal{A}}(u) \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(u + h_{0,1}^{u\pm}) dx \right) \right) \\ & \leq \sup_{u \in B(u_k, \tilde{\delta})} \exp \left( -\frac{\ell}{\varepsilon} \int_0^1 V(u + h_{0,1}^{u\pm}) dx \right) \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(\overline{B(u_k, \tilde{\delta})}). \end{aligned} \quad (6.27)$$

Due to the lower semi-continuity of  $I_{0,1}$ , we can choose  $\tilde{u}_k \in B(u_k, \tilde{\delta})$  so that

$$I_{0,1}(\tilde{u}_k) \leq \inf_{u \in B(u_k, \tilde{\delta})} I_{0,1}(u) + \gamma. \quad (6.28)$$

Then the first factor in (6.27) can be bounded above by

$$\exp \left( -\frac{1}{\varepsilon} \left( \ell \int_0^1 V(\tilde{u}_k + h_{0,1}^{u\pm}) dx - 2\tilde{\delta} \ell_+ \sup_{|v| \leq \|\tilde{u}_k\|_\infty + M + 1} |V'(v)| \right) \right), \quad (6.29)$$

and we need a bound on  $\|\tilde{u}_k\|_\infty$ . First we recall that  $u_k \in \mathcal{K}_{\ell_+ R}$ , which by definition gives  $I_{0,1}(u_k) \leq \ell_+ R$ . Together with the definition of  $\tilde{u}_k$ , this gives

$$I_{0,1}(\tilde{u}_k) \stackrel{(6.28)}{\leq} \ell_+ R + \gamma.$$

Recalling the homogeneous boundary conditions, this implies that

$$\|\tilde{u}_k\|_\infty \leq \frac{1}{2} \int_0^1 |\partial_x \tilde{u}_k| dx \leq \frac{1}{2} \left( \int_0^1 |\partial_x \tilde{u}_k|^2 dx \right)^{1/2} \leq \sqrt{2^{-1}(\ell_+ R + 1)}.$$

Hence, the definition (6.22) of  $\tilde{\delta}$  implies that the bound in (6.29) improves to

$$\exp \left( -\frac{1}{\varepsilon} \left( \ell \int_0^1 V(\tilde{u}_k + h_{0,1}^{u\pm}) dx - 2\gamma \right) \right). \quad (6.30)$$

On the other hand, the Gaussian large deviation bound (6.16) and the definition (6.28) of  $\tilde{u}_k$  imply that for every  $k$  there exists an  $\varepsilon_0 > 0$  such that for  $\ell_+ \varepsilon \leq \varepsilon_0$  we have

$$\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(\overline{B(u_k, \tilde{\delta})}) \leq \exp \left( -\frac{1}{\ell\varepsilon} (I_{0,1}(\tilde{u}_k) - 2\gamma) \right). \quad (6.31)$$

As there are only finitely many  $u_k$  (the selection of which does not depend on  $\mathcal{A}$ ), we can find an  $\varepsilon_0$  such that this bound holds for all  $\tilde{u}_k$  simultaneously and such that (6.25) holds as well.

Substituting (6.30) and (6.31) into (6.27) gives for each  $k$  that

$$\begin{aligned} & \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon,0,0}} \left( \mathbf{1}_{B(u_k, \tilde{\delta}) \cap \mathcal{A}} \exp \left( -\frac{1}{\varepsilon} \ell \int_0^1 V(u + h_{0,1}^{u\pm}) dx \right) \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} \left( E_\ell^{u\pm}(\tilde{u}_k) - \left( 2 + \frac{2}{\ell_-} \right) \gamma \right) \right) \\ & \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(u_k, \tilde{\delta})} E_\ell^{u\pm}(u) - \left( 2 + \frac{2}{\ell_-} \right) \gamma \right) \right) \\ & \stackrel{(6.22)}{\leq} \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(u_k, \tilde{\delta})} E_\ell^{u\pm}(u) - \left( 2 + \frac{2}{\ell_-} \right) \gamma \right) \right). \end{aligned}$$

After relabelling  $\gamma$  (for instance by a factor of 6), the above bound together with (6.23), (6.25), and (6.26) finishes the proof of (6.21).  $\square$

The proof of the lower bound (6.15) relies on the classical Cameron-Martin Theorem. In the current context it can be stated as follows.

**Theorem 6.5** (Cameron-Martin Thm. e.g. [Hai09, Thm 3.41]). *For a fixed  $f \in C([x_-, x_+])$ , define the shift map  $T_f: C([x_-, x_+]) \rightarrow C([x_-, x_+])$  by  $T_f(u) = u + f$ . Then the image measure  $T_f^* \mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  is absolutely continuous with respect to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}$  if and only if  $f \in H_0^1(x_-, x_+)$ . In that case the Radon-Nykodym derivative is given by*

$$\frac{dT_f^* \mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}}{d\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0}}(u) = \exp \left( -\frac{1}{\varepsilon} I_{x_-, x_+}(f) + \frac{1}{\varepsilon} \int_{x_-}^{x_+} \partial_x f(x) du(x) \right). \quad (6.32)$$

Here, as in the case of Brownian motion, the stochastic integral term  $\frac{1}{\varepsilon} \int_{x_-}^{x_+} \partial_x f(x) du(x)$  can be defined as the limit of Riemann sums in  $L^2(\mathcal{W}_{\varepsilon, (x_-, x_+)}^{0,0})$ . In particular, it is a linear mapping in  $u$  defined for all  $u$  in a measurable subspace of  $C([x_-, x_+])$  of full measure (See e.g. [Hai09, Sec. 3]).

Note that (6.32) can formally be derived by expanding the square in the non-rigorous expression (3.3).

*Proof of Lemma 6.3.* We can assume that  $u_* \in H^1$ , because otherwise the bound is trivial. As in the proof of the upper bound, (6.15) only gets stronger when we take a smaller  $\delta$ . Therefore, it is sufficient to show (6.15) with  $\delta$  replaced by

$$\tilde{\delta} := \gamma \left( \sup_{|v| \leq M+1} |V'(v)| \ell_+ \right)^{-1} \wedge \delta \wedge 1. \quad (6.33)$$

We begin by stating the simplistic bound

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}} \left( \mathbf{1}_{B(u_*, \tilde{\delta})}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \exp \left( -\frac{1}{\varepsilon} \sup_{u \in B(u_*, \tilde{\delta})} \int_{x_-}^{x_+} V(u) dx \right) \mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(B(u_*, \tilde{\delta})). \end{aligned} \quad (6.34)$$

Due to the assumption (6.14) on  $u_*$  and the definition (6.33) of  $\tilde{\delta}$ , we get that

$$\begin{aligned} \sup_{u \in B(u_*, \tilde{\delta})} \int_{x_-}^{x_+} V(u) dx &\leq \int_{x_-}^{x_+} V(u_*) dx + \sup_{|v| \leq M+1} |V'(v)| \tilde{\delta} \ell_+ \\ &\leq \int_{x_-}^{x_+} V(u_*) dx + \gamma. \end{aligned}$$

It only remains to derive a lower bound on  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(B(u_*, \tilde{\delta}))$  in terms of the Gaussian energy. To this end, we again transform  $u_*$  to an interval of length one and shift it in a way that it satisfies homogenous boundary conditions, as in the proof of Lemma 6.2. To be more precise, we assume that  $u$  is distributed according to  $\mathcal{W}_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$  and apply the affine transformation  $T$  defined in (6.17). Then  $Tu = \hat{u}$  is distributed according to  $\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}$ . Therefore, we have to bound the probability

$\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(B(\hat{u}_*, \tilde{\delta}))$ , where  $\hat{u}_* := Tu_*$ . This can be obtained using the Cameron-Martin Theorem 6.5 with  $f := \hat{u}_*$ . According to (6.32), we have

$$\begin{aligned} \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(B(\hat{u}_*, \tilde{\delta})) &= \exp\left(-\frac{1}{\ell\varepsilon}I_{0,1}(\hat{u}_*)\right) \\ &\quad \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}}\left(\mathbf{1}_{B(0,\tilde{\delta})}(\hat{u}) \exp\left(\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right)\right). \end{aligned}$$

Now we will use the trick of sneaking in a cosh function. To this end, we remark that the map  $\hat{u} \mapsto \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)$  is linear in  $\hat{u}$ . Also, the measure  $\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}$  is invariant under the mapping  $\hat{u} \mapsto -\hat{u}$  and this mapping leaves the ball  $B(0, \tilde{\delta})$  invariant. Hence, the last expectation is equal to

$$\mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}}\left(\mathbf{1}_{B(0,\tilde{\delta})}(\hat{u}) \exp\left(-\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right)\right).$$

Therefore, we can write

$$\begin{aligned} &\mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}}\left(\mathbf{1}_{B(0,\tilde{\delta})}(\hat{u}) \exp\left(\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right)\right) \\ &= \frac{1}{2} \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}}\left(\mathbf{1}_{B(0,\tilde{\delta})}(\hat{u}) \left[ \exp\left(\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right) \right. \right. \\ &\quad \left. \left. + \exp\left(-\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right) \right] \right) \\ &= \mathbb{E}_{(0,1)}^{\mathcal{W}_{\ell\varepsilon, 0,0}}\left(\mathbf{1}_{B(0,\tilde{\delta})}(\hat{u}) \cosh\left(\frac{1}{\ell\varepsilon} \int_0^1 \partial_x \hat{u}_*(x) d\hat{u}(x)\right)\right) \\ &\geq \mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(B(0, \tilde{\delta})). \end{aligned}$$

We claim that there exists an  $\varepsilon_0 > 0$  such that for all  $\ell \leq \ell_+$  and all  $\varepsilon \leq \varepsilon_0$  this probability is larger than  $\exp(-\varepsilon^{-1}\gamma)$ . Actually, (6.16) even implies that for any  $\tilde{\gamma} > 0$  there exists  $\tilde{\varepsilon}_0 > 0$  such that, for  $\ell\varepsilon \leq \tilde{\varepsilon}_0$ , we have the stronger bound

$$\mathcal{W}_{\ell\varepsilon, (0,1)}^{0,0}(\mathbb{C}B(0, \tilde{\delta})) \leq \exp\left(-\frac{1}{\ell\varepsilon} \left(\inf_{\hat{u} \in \mathbb{C}B(0, \tilde{\delta})} I_{0,1}(\hat{u}) - \tilde{\gamma}\right)\right).$$

Note that this  $\varepsilon_0$  also depends on  $\sup_{|v| \leq M+1} |V'(v)|$  as we have potentially decreased  $\delta$  in the first step. Then in order to conclude, it is sufficient to observe that

$$\frac{1}{\ell\varepsilon} I_{0,1}(\hat{u}_*) = \frac{1}{\varepsilon} (I_{x_-, x_+}(u_*) - I_{x_{\pm}}^{u_{\pm}}).$$

□

Now the proofs of Propositions 3.4 and 3.5 are straightforward. We begin with the upper bound, Proposition 3.4.

*Proof of Proposition 3.4.* We want to derive a bound on

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(\mathcal{A}) = \frac{\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}}\left(\mathbf{1}_{\mathcal{A}}(u) \exp\left(-\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx\right)\right)}{\mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_{\varepsilon, u_-, u_+}}\left(\exp\left(-\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx\right)\right)}. \quad (6.35)$$



The assumptions on  $\mathcal{A}$  in Proposition 3.4 are identical to those in Lemma 6.2, so we can conclude from (6.13) that

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in B(\mathcal{A}, \delta)} E(u) - I_{x_\pm}^{u_\pm} - \gamma \right) \right) \end{aligned}$$

for  $\varepsilon \leq \varepsilon_0$ . Also this  $\varepsilon_0$  depends on  $M, R, \ell_+, \delta$ , and  $\gamma$  but not on the particular choice of  $x_\pm, u_\pm$ . It only depends on  $\mathcal{A}$  through the condition (6.12) and on  $V$  through the local Lipschitz constant.

To get a lower bound on the denominator in (6.35), we observe that for every set of boundary conditions  $u_\pm$ , there exists at least one minimizer  $u_*$  of  $E$  given these boundary conditions. Furthermore, this minimizer attains only values in  $[-M, M]$ . This is clear because replacing  $u_*$  by  $u_* \wedge M \vee (-M)$  only decreases the energy. Therefore, for any  $\delta > 0$ , we get from (6.15) that

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \mathbf{1}_{B(u_*, \delta)}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \exp \left( -\frac{1}{\varepsilon} (E(u_*) - I_{x_\pm}^{u_\pm} + \gamma) \right) \end{aligned}$$

for  $\varepsilon \leq \varepsilon_0$ , where  $\varepsilon_0$  satisfies the same uniformity assumptions as above. This finishes the argument.  $\square$

The proof of the lower bound is similar.

*Proof of Proposition 3.5.* To derive a lower bound on  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(\mathcal{A})$  for a given  $\gamma$  we choose  $u_\gamma$  as in (3.19). Then we can write using (6.15)

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \mathbf{1}_{B(\mathcal{A}, \delta)}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \mathbf{1}_{B(u_\gamma, \delta)}(u) \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \geq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in \mathcal{A}} E(u) - I_{x_\pm}^{u_\pm} + 2\gamma \right) \right), \end{aligned}$$

for  $\varepsilon \leq \varepsilon_0$  where  $\varepsilon_0$  can again be chosen uniformly.

To derive a uniform upper bound on the normalization constant we only need to observe that for any  $M < \infty$  there exists an  $R < \infty$  such that for all  $u_\pm \in [-M, M]$ , we have

$$\inf_{u \in \mathcal{A}^{\text{bc}}} E(u) \leq R.$$

Then (6.13) implies that there exists  $\varepsilon_0 > 0$  such that uniformly for  $\varepsilon \leq \varepsilon_0$

$$\begin{aligned} \mathbb{E}_{(x_-, x_+)}^{\mathcal{W}_\varepsilon, u_-, u_+} \left( \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u) dx \right) \right) \\ \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in \mathcal{A}^{\text{bc}}} E(u) - I_{x_\pm}^{u_\pm} + \gamma \right) \right). \end{aligned}$$

This establishes (3.18).

□

#### 6.4. Proof of the one-point distribution lemma.

*Proof of Lemma 4.1.* First we remark that, heuristically, the “most difficult” point to consider is  $x_0 = 0$ . We present the following proof for precisely this case. The same proof carries over for any point  $x_0$  (with only trivial modifications), but we present it for  $x_0 = 0$  since it simplifies the notation slightly and makes the main ideas stand out.

Also notice that by the symmetry of the potential (cf. Assumption 1.1) and the representation (1.5), it suffices to prove

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} \left( u(x_0) \geq M \right) \leq \exp \left( -\frac{M}{\varepsilon C_2} \right).$$

In fact, it will be convenient to establish the estimate in the form

$$\mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} \left( u(0) \geq 4M \right) \leq \exp \left( -\frac{M}{\varepsilon \tilde{C}_2} \right), \quad (6.36)$$

which is of course equivalent for  $C_2 := 4\tilde{C}_2$ . Thus consider the set of functions

$$\mathcal{A} := \{u \in C([-L_\varepsilon, L_\varepsilon]) : u(-L_\varepsilon) = -1, u(L_\varepsilon) = 1, \text{ and } u(0) \geq 4M\}. \quad (6.37)$$

Define  $x_\pm^{3M}$  as follows:

$$x_-^{3M} := \sup\{x \leq 0 : u(x) \leq 3M\} \quad \text{and} \quad x_+^{3M} := \inf\{x \geq 0 : u(x) \leq 3M\}.$$

Notice that we may assume without loss of generality that  $M \geq 1$ , and hence, because of the boundary conditions  $u(-L_\varepsilon) = -1$  and  $u(L_\varepsilon) = 1$ , the points  $x_-^{3M} < 0 < x_+^{3M}$  are well-defined for every  $u \in \mathcal{A}$ . The set  $\mathcal{A}$  can then be divided into the following two sets:

$$\mathcal{A}_1 := \{u \in \mathcal{A} : \max\{|x_-^{3M}|, x_+^{3M}\} > 1\},$$

$$\mathcal{A}_2 := \{u \in \mathcal{A} : \max\{|x_-^{3M}|, x_+^{3M}\} \leq 1\}.$$

To bound the probability of  $\mathcal{A}_1$ , we will use bounds on the potential and a reflection argument. For  $\mathcal{A}_2$  we will use a rescaling argument and the large deviation bound (3.17). The two cases are illustrated in Figure 6.1.

Step 1. We treat  $\mathcal{A}_1$  first. For  $u \in \mathcal{A}_1$  we have

$$x_+^{3M} - x_-^{3M} \geq 1. \quad (6.38)$$

The idea is to introduce a reflection over the line  $u = 2M$  that preserves the Gaussian measure, and use the decrease of the energy (1.6) under this reflection.

We begin by collecting some facts about the potential  $V$ . To begin with, according to the growth estimate in (1.3),  $V$  grows superlinearly at infinity. Hence, we may choose  $C_3$  sufficiently large so that the following two properties are satisfied. On the one hand,  $V$  grows at least linearly on  $[C_3, \infty)$ , i.e., there exists  $C_4 < \infty$  such that for  $u_1 \geq u_2 \geq C_3$ , there holds

$$V(u_1) - V(u_2) \geq 1/C_4 (u_1 - u_2). \quad (6.39)$$

On the other hand,  $V(C_3) \geq V(0)$ , so that in particular

$$V(C_3) = \sup_{u \in [0, C_3]} V(u). \quad (6.40)$$

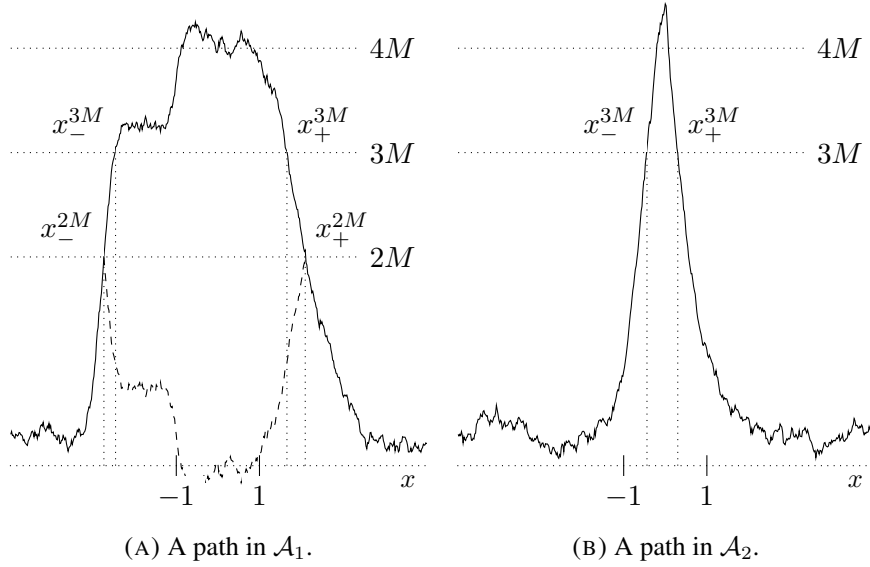


FIGURE 6.1. The two different cases. To show that  $\mathcal{A}_1$  has small probability, we reflect between the  $x_{\pm}^{2M}$ . This decreases the *potential energy*. The probability of  $\mathcal{A}_2$  can be bounded using a large deviation argument.

We will use the fact that (6.39) and (6.40) together imply that as long as  $u_1 \geq C_3$ , then

$$u_1 \geq |u_2| \Rightarrow V(u_1) \geq V(|u_2|). \quad (6.41)$$

Now we are ready to reflect. Define  $x_{\pm}^{2M}$  analogously to  $x_{\pm}^{3M}$  (noting as above that they are well-defined for paths in the set of interest). Consider the reflection operator  $R_{x_{-}^{2M}}^{x_{+}^{2M}}$  defined as

$$R_{x_{-}^{2M}}^{x_{+}^{2M}} u(x) := \begin{cases} u(x) & \text{if } x \notin (x_{-}^{2M}, x_{+}^{2M}) \\ 4M - u(x) & \text{if } x \in (x_{-}^{2M}, x_{+}^{2M}) \end{cases}, \quad (6.42)$$

which for the purposes of this lemma we will abbreviate with  $R$ . In order to have  $R$  well-defined for all continuous paths  $u$ , we define it to be the identity for those paths  $u$  that never exceed the level  $2M$ .

Notice that  $x_{-}^{2M}$  is a *right* but not a *left* stopping point, and similarly  $x_{+}^{2M}$  is a *left* but not a *right* stopping point. In particular, the strong Markov property (3.9) does not directly imply that  $R$  leaves  $\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$  invariant. Indeed, it is not true that under  $\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$  the conditional distribution of  $u(x)$  for  $x \in [x_{-}^{2M}, x_{+}^{2M}]$ , given the path outside of this interval, is a Brownian bridge.

Still, it is true that the reflection operator  $R$  preserves  $\mathcal{W}_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}$ . To see this, introduce auxiliary stopping points

$$\chi_{-}^{2M} := \inf\{x \geq -L_{\varepsilon} : u(x) = 2M\} \quad \text{and} \quad \chi_{+}^{2M} := \sup\{x \leq L_{\varepsilon} : u(x) = 2M\}.$$

As above in (4.18), we use the convention that  $\chi_{\pm}^{2M} = \mp L_{\varepsilon}$  if these sets are empty.

On  $\mathcal{A}$ , these points are well-defined and we automatically have  $[x_{-}^{2M}, x_{+}^{2M}] \subseteq [\chi_{-}^{2M}, \chi_{+}^{2M}]$ . The points  $\chi_{\pm}^{2M}$  are left and right stopping points. Therefore, (3.9)

implies that the reflection operators  $R_{\chi_-^{2M}}^{\chi_+^{2M}}$ ,  $R_{x_+^{2M}}^{\chi_+^{2M}}$ , and  $R_{\chi_-^{2M}}^{x_-^{2M}}$  (defined in the same way as  $R$ ) preserve  $\mathcal{W}_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}$ . Observing that

$$R = R_{\chi_-^{2M}}^{\chi_+^{2M}} \circ R_{x_+^{2M}}^{\chi_+^{2M}} \circ R_{\chi_-^{2M}}^{x_-^{2M}},$$

we conclude that  $R$  also preserves  $\mathcal{W}_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1}$ .

We now develop a quantitative, pointwise estimate of the effect of  $R$  on the “bulk energy”  $V(u)$ . By the definition of  $x_\pm^{2M}$ , we have that  $u(x) \geq 2M$  for all  $x \in [x_-^{2M}, x_+^{2M}]$ , the set where  $R$  acts. Hence, it suffices to consider the effect of  $R$  when  $u(x) \geq 3M$  and when  $u(x) \in [2M, 3M)$ . We will first establish that on the set

$$\{x \in [x_-^{2M}, x_+^{2M}]: u(x) \geq 3M\},$$

$R$  decreases the bulk energy significantly. Indeed, on this set,  $|Ru| \leq u - 2M$  and  $u - 2M \geq M \geq C_3$ , so that

$$V(|Ru|) \stackrel{(6.41)}{\leq} V(u - 2M), \quad (6.43)$$

which together with (6.39) implies that for  $u(x) \geq 3M$ ,

$$\begin{aligned} V(u(x)) - V(Ru(x)) &\stackrel{(1.3)}{=} V(u(x)) - V(|Ru(x)|) \\ &\stackrel{(6.43)}{\geq} V(u(x)) - V(u(x) - 2M) \\ &\stackrel{(6.39)}{\geq} 2M/C_4, \end{aligned} \quad (6.44)$$

which holds in particular on all of  $[x_-^{3M}, x_+^{3M}]$ . On the other hand, if instead  $u(x) \in [2M, 3M)$ , then the bulk energy still decreases under  $R$ . Indeed, we have for  $u(x) \in [2M, 3M)$  that  $u(x) \geq Ru(x) \geq M$ , so that by (6.41) we know

$$V(u(x)) - V(Ru(x)) \geq 0. \quad (6.45)$$

Combining (6.44) and (6.45) implies that for all  $u \in \mathcal{A}_1$ , we have

$$\begin{aligned} &\int_{(-L_\varepsilon, L_\varepsilon)} (V(u) - V(Ru)) \, dx \\ &= \int_{(x_-^{2M}, x_+^{2M})} (V(u) - V(Ru)) \, dx \\ &\geq 2M(x_+^{3M} - x_-^{3M})/C_4 \stackrel{(6.38)}{\geq} 2M/C_4. \end{aligned} \quad (6.46)$$

We are now ready to estimate the probability of  $\mathcal{A}_1$ . Indeed, we have that

$$\begin{aligned}
1 &\geq \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{R}\mathcal{A}_1) \\
&\stackrel{(1.5)}{=} \frac{1}{\mathcal{Z}} \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mathcal{W}_\varepsilon,-1,1} \left[ \mathbf{1}_{\mathcal{R}\mathcal{A}_1}(u) \exp \left( -\frac{1}{\varepsilon} \int V(u) dx \right) \right] \\
&\stackrel{\text{inv. of R}}{=} \frac{1}{\mathcal{Z}} \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mathcal{W}_\varepsilon,-1,1} \left[ \mathbf{1}_{\mathcal{A}_1}(u) \exp \left( -\frac{1}{\varepsilon} \int V(\mathcal{R}u) dx \right) \right] \\
&= \frac{1}{\mathcal{Z}} \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mathcal{W}_\varepsilon,-1,1} \left[ \mathbf{1}_{\mathcal{A}_1}(u) \exp \left( -\frac{1}{\varepsilon} \int (V(\mathcal{R}u) - V(u)) dx \right. \right. \\
&\quad \left. \left. -\frac{1}{\varepsilon} \int V(u) dx \right) \right] \\
&\stackrel{(6.46)}{\geq} \exp \left( \frac{2M}{C_4 \varepsilon} \right) \frac{1}{\mathcal{Z}} \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mathcal{W}_\varepsilon,-1,1} \left[ \mathbf{1}_{\mathcal{A}_1}(u) \exp \left( -\frac{1}{\varepsilon} \int V(u) dx \right) \right] \\
&= \exp \left( \frac{2M}{C_4 \varepsilon} \right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1),
\end{aligned}$$

where  $\mathcal{Z} = \mathcal{Z}_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  is the normalization constant for  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$  and all of the integrals are over  $[-L_\varepsilon, L_\varepsilon]$ . Moving the exponential to the other side of the inequality, we get that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_1) \leq \exp \left( -\frac{2M}{C_4 \varepsilon} \right), \quad (6.47)$$

which gives (6.36) for  $\mathcal{A}_1$  with  $\tilde{C}_2 = C_4/2$ .

**Step 2.** Now consider the set  $\mathcal{A}_2$ . Here we will use a rescaling argument and the large deviation bound (3.17). For  $u \in \mathcal{A}_2$ , we can define

$$\chi_- := \inf\{x \in [-1, 0] : u(x) = 3M\}, \quad \chi_+ := \sup\{x \in [0, 1] : u(x) = 3M\},$$

with the understanding that  $\chi_\pm = 0$  if these sets are empty. These random variables are left and right stopping points. Hence, the strong Markov property (3.10) implies that

$$\begin{aligned}
&\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_2) \\
&= \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u(0) \geq 4M \text{ and } \chi_\pm \neq 0 \right) \\
&= \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mu_\varepsilon,-1,1} \left( \mu_{\varepsilon,(\chi_-, \chi_+)}^{3M,3M} (u(0) \geq 4M) \mathbf{1}_{\{\chi_\pm \neq 0\}} \right). \quad (6.48)
\end{aligned}$$

Therefore, if we can show that

$$\mu_{\varepsilon,(x_-, x_+)}^{3M,3M} \left( u(0) \geq 4M \right) \leq \exp \left( -\frac{M}{\tilde{C}_2 \varepsilon} \right) \quad (6.49)$$

for all  $\varepsilon$  sufficiently small (uniformly for  $-x_-, x_+ \in (0, 1]$ ), then the combination of (6.48) and (6.49) concludes the proof of (6.36). We can see (6.49) by rescaling. Indeed, if we transform  $(x_-, x_+)$  into  $[-1, 1]$  by applying the affine change of

variables  $x \rightarrow \frac{\Delta x}{2} x + \frac{x_- + x_+}{2}$  where  $\Delta x := x_+ - x_-$ , we see that

$$\begin{aligned} & \mu_{\varepsilon, (x_-, x_+)}^{3M, 3M} \left( u(0) \geq 4M \right) \\ &= \frac{1}{\mathcal{Z}} \mathbb{E}_{\mathcal{W}_{\varepsilon, (x_-, x_+)}^{3M, 3M}} \left( \mathbf{1}_{\{u(0) > 4M\}} \exp \left( -\frac{1}{\varepsilon} \int_{x_-}^{x_+} V(u(x)) dx \right) \right) \\ &= \frac{1}{\mathcal{Z}} \mathbb{E}_{\mathcal{W}_{\tilde{\varepsilon}, (-1, 1)}^{3M, 3M}} \left( \mathbf{1}_{\{\hat{u}(\frac{x_- + x_+}{2}) > 4M\}} \exp \left( -\frac{(\Delta x)^2}{4\tilde{\varepsilon}} \int_{-1}^1 V(\hat{u}(x)) dx \right) \right), \end{aligned} \quad (6.50)$$

where  $\mathcal{Z} = \mathcal{Z}_{\varepsilon, (x_-, x_+)}^{3M, 3M}$  is the normalization constant for  $\mu_{\varepsilon, (x_-, x_+)}^{3M, 3M}$  and  $\tilde{\varepsilon} := \frac{1}{2}\varepsilon\Delta x$ . Now, we observe that the family of potentials

$$\left\{ \frac{(\Delta x)^2}{4} V : 0 < \Delta x \leq 2 \right\}$$

is locally uniformly Lipschitz. In particular, applying Proposition 3.4 for  $\gamma$  and  $\delta$  fixed to say  $\gamma = \delta = 1$ , there exists  $\varepsilon_0 > 0$  such that, for  $\tilde{\varepsilon} \leq \varepsilon_0$  and uniformly in  $x_{\pm}$ , we have

$$\mu_{\tilde{\varepsilon}, (-1, 1)}^{3M, 3M}(\mathcal{B}_+^{\text{bc}}) \leq \exp \left( -\frac{1}{\tilde{\varepsilon}} \left( \inf_{\hat{u} \in B(\mathcal{B}_+^{\text{bc}}, 1)} E_{\Delta x}(\hat{u}) - \inf_{\hat{u} \in \mathcal{B}^{\text{bc}}} E_{\Delta x}(\hat{u}) - 1 \right) \right).$$

Note that the choice of  $\varepsilon_0$  depends on  $M$ .

Here we use the notation

$$E_{\Delta x}(\hat{u}) := \int_{-1}^1 \left( \frac{1}{2} (\partial_x \hat{u})^2 + \frac{(\Delta x)^2}{4} V(\hat{u}) \right) dx$$

and

$$\begin{aligned} \mathcal{B}^{\text{bc}} &:= \{\hat{u} \in C([-1, 1]) : \hat{u}(\pm 1) = 3M\}, \\ \mathcal{B}_+^{\text{bc}} &:= \{\hat{u} \in C([-1, 1]) : \hat{u}(\pm 1) = 3M, \hat{u}((x_- + x_+)/2) \geq 4M\}. \end{aligned}$$

Hence, as  $\tilde{\varepsilon} \leq \varepsilon$ , to establish (6.49) it will be sufficient for us to show

$$\inf_{B(\mathcal{B}_+^{\text{bc}}, 1)} E_{\Delta x} - \inf_{\mathcal{B}^{\text{bc}}} E_{\Delta x} \geq \frac{M}{\tilde{C}_2},$$

and we will in fact establish the stronger bound

$$\inf_{B(\mathcal{B}_+^{\text{bc}}, 1)} E_{\Delta x} - \inf_{\mathcal{B}^{\text{bc}}} E_{\Delta x} \geq (M-1)^2 \stackrel{M \geq 4}{\geq} \frac{M^2}{2}.$$

We will establish the first inequality by way of a variational argument. Notice that we may assume that the infima are achieved (if not, a simple approximation argument suffices), and so let

$$\hat{u}_1 := \operatorname{argmin}_{B(\mathcal{B}_+^{\text{bc}}, 1)} E_{\Delta x}, \quad \hat{u}_2 := \operatorname{argmin}_{\mathcal{B}^{\text{bc}}} E_{\Delta x}.$$

Observe that automatically  $\hat{u}_1((x_- + x_+)/2) \geq 4M - 1$ .

We define the auxiliary function  $\hat{u}_3 := \min\{\hat{u}_1, 3M\}$ . Notice that according to the growth assumption (1.3) (or see (6.39)):

$$V(\hat{u}_1(x)) \geq V(3M) \quad \text{on } \{\hat{u}_1 \geq 3M\}. \quad (6.51)$$

On the other hand, since  $\hat{u}_3 \in \mathcal{B}^{\text{bc}}$  and as  $\hat{u}_2$  is the minimizer over  $\mathcal{B}^{\text{bc}}$ , we have

$$\begin{aligned}
E_{\Delta x}(\hat{u}_1) &= E_{\Delta x}(\hat{u}_2) \\
&\geq E_{\Delta x}(\hat{u}_1) - E_{\Delta x}(\hat{u}_3) \\
&= \int_{\{u \geq 3M\}} \left( (\partial_x \hat{u}_1)^2 + \frac{(\Delta x)^2}{4} (V(\hat{u}_1) - V(3M)) \right) dx \\
&\stackrel{(6.51)}{\geq} \int_{-1}^1 (\partial_x \max\{\hat{u}_1 - 3M, 0\})^2 dx \\
&\geq \left( \sup_{x \in [-1, 1]} \max\{\hat{u}_1 - 3M, 0\} \right)^2 \\
&\geq (M - 1)^2.
\end{aligned}$$

This concludes the proof of (6.36) for  $\mathcal{A}_2$  and establishes the lemma.  $\square$

### 6.5. Proofs of lemmas from the lower bound of Theorem 1.5.

*Proof of Lemma 4.3.* Since the proof is similar to (and simpler than) the proof of the upper bound in Theorem 1.5, we will be somewhat brief. Our goal is to bound above by  $2/3$  the complementary event, namely that  $u(x) > 0$  for some  $x \in [-L_\varepsilon, -2\ell]$  or that  $|u(x_k)| > M$  for some  $k$  in the index set. As in the proof of the upper bound, the probability that  $|u(x_k)| > M$  can be shown to be exponentially small in  $M/\varepsilon$ , cf. (4.9). It remains to bound above the probability that  $u(x) > 0$  for some  $x \in [-L_\varepsilon, -2\ell]$  and  $|u(x_k)| \leq M$  for all  $k$ .

Now fix  $\delta > 0$  sufficiently small so that the estimates from the upper bound of Theorem 1.5 apply. The set  $\{u \in \mathcal{C}\mathcal{A}_1 : u(x) > 0 \text{ for some } x \in [-L_\varepsilon, -2\ell]\}$  is contained within the union of:

- (1) functions with more than one  $\delta^-$  layer (exponentially unlikely by the upper bound of Theorem 1.5),
- (2) functions with a  $\delta^-$  layer longer than  $2\ell$  (exponentially unlikely for  $\delta^2\ell$  large, according to the calculation in Step 3 of the proof of the upper bound, cf. (4.13)),
- (3) functions with one and only one  $\delta^-$  layer, which is at most length  $2\ell$  and is contained in  $[-L_\varepsilon, 0]$ ,
- (4) functions with one and only one  $\delta^-$  layer, which is at most length  $2\ell$  and is contained in  $[-2\ell, L_\varepsilon]$ , and such that  $u(x) > 0$  for some  $x \in [-L_\varepsilon, -2\ell]$ .

By symmetry properties of the measure, i.e. the symmetry with respect to point reflection of the graph at  $x = 0$  and  $u = 0$ , the probability of a  $\delta^-$  layer contained in  $[-L_\varepsilon, 0]$  is equal to the probability of a  $\delta^-$  layer contained in  $[0, L_\varepsilon]$ , hence neither can be more than  $1/2$ . Therefore, the probability of the event described in point (3) is less than or equal to  $1/2$ .

By the calculations referred to above, the sum of the probabilities of the sets described in (1)-(3) is bounded by  $1/2$  plus exponentially small terms, so we are finished if we can show that the probability of the set described in (4) is also exponentially small, namely, the probability that:  $u(x) > 0$  for some  $x \in [-L_\varepsilon, -2\ell]$ ,  $|u(x_k)| \leq M$  for all  $k$ , and there is one and only one  $\delta^-$  layer, which is at most  $2\ell$  and is contained in  $[-2\ell, L_\varepsilon]$ . Note that the latter implies that  $u \leq 1 - \delta$  on  $[-L_\varepsilon, -2\ell]$ .

This bound is easy to obtain by breaking into subintervals (using conditioning) and using the large deviation estimate (3.17). Indeed, we reduce to probabilities of

the form

$$\mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_{k-2}, u_{k+2}} \left( u \leq 1 - \delta \text{ and } u(\hat{x}) = 0 \text{ for some } \hat{x} \in (x_{k-1}, x_{k+1}) \right),$$

where  $u_{k-2}$  and  $u_{k+2}$  are arbitrary boundary values in  $[-M, 1 - \delta]$  and  $k \in \{-(N_\varepsilon - 2), -(N_\varepsilon - 3), \dots, -3\}$ . (We also need to consider the boundary interval, where  $\hat{x} \in (x_{-N_\varepsilon}, x_{-(N_\varepsilon-2)})$ . As usual, this is no more difficult than the bound for the interior intervals.) After applying Proposition 3.4 (with  $\tilde{\delta} = \delta/2$ ), it remains only to introduce an energetic bound. The bound from Lemma 6.6 below suffices.

Before stating the energy lemma, we explain the idea in words: If we take a  $\delta/2$  ball around the set of interest, then on  $[x_{k-1}, x_{k+1}]$ , there is a point  $x_0$  such that  $u(x_0) \geq -\delta/2$ . For  $\ell$  large, the energy minimizer needs to come very close to  $\pm 1$  someplace in  $[x_{k-2}, x_{k-1}]$  and  $[x_{k+1}, x_{k+2}]$ , (say within  $\delta/4$ ), and since it cannot come this close to  $+1$ , it is forced into a small neighborhood of  $-1$ . Consequently, the large excursion from  $-1$  at  $x_0$  costs almost  $c_0$  energy. We give the precise statement below and prove the lemma at the end of the subsection.

**Lemma 6.6.** *There exists  $C < \infty$  with the following property. For any  $M$  large enough and  $\delta > 0$  small enough, consider the boundary conditions  $u_\pm \in [-M, 1 - \delta]$  and define the sets*

$$\begin{aligned} \mathcal{A}^{\text{bc}} &:= \{u \in C([-2\ell, 2\ell]) : u(-2\ell) = u_- \text{ and } u(2\ell) = u_+\}, \\ \mathcal{A}_0^{\text{bc}} &:= \{u \in \mathcal{A}^{\text{bc}} : u(x) \leq 1 - \delta/2 \text{ for all } x \in [-2\ell, 2\ell] \text{ and} \\ &\quad \text{there is an } x_0 \in [-\ell, \ell] \text{ such that } u(x_0) \geq -\delta/2\}. \end{aligned}$$

Then there exists  $\ell_0 = \ell_0(M, \delta)$  such that for  $\ell \geq \ell_0$  there holds

$$\inf_{u \in \mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) - \inf_{u \in \mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \geq c_0 - C\delta.$$

Proposition 3.4 and Lemma 6.6 together give

$$\begin{aligned} \mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_{k-2}, u_{k+2}} \left( u \leq 1 - \delta \text{ and } u(x_0) > 0 \text{ for some } x_0 \in (x_{k-1}, x_{k+1}) \right) \\ \leq \exp \left( -\frac{c_0 - C\delta - \gamma}{\varepsilon} \right). \end{aligned}$$

Finally, we now choose  $\gamma$  and  $\delta$  sufficiently small and sum over the order  $N_\varepsilon \sim L_\varepsilon$  intervals. Bearing in mind the bound (1.10) on  $L_\varepsilon$ , we observe that there is also an exponentially small probability of the final set that we have studied.  $\square$

*Proof of Lemma 6.6.* We will be brief, since the proof is similar to the proof of Lemma 2.5.

First of all, fix  $M$  large and  $\delta$  small. The infimum of the energy over  $\mathcal{A}^{\text{bc}}$  is less than or equal to the minimum of the energy over functions with  $u(\pm 2\ell) = u_\pm$  and  $u(0) = -1$ . By a standard construction, we have

$$\inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \leq \varphi_{-1}(u_-) + \varphi_{-1}(u_+) + o(1)_{\ell \uparrow \infty}.$$

In particular, for  $\ell_0$  large enough and  $\ell \geq \ell_0$ , one has

$$\inf_{\mathcal{A}^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \leq \varphi_{-1}(u_-) + \varphi_{-1}(u_+) + \delta. \quad (6.52)$$



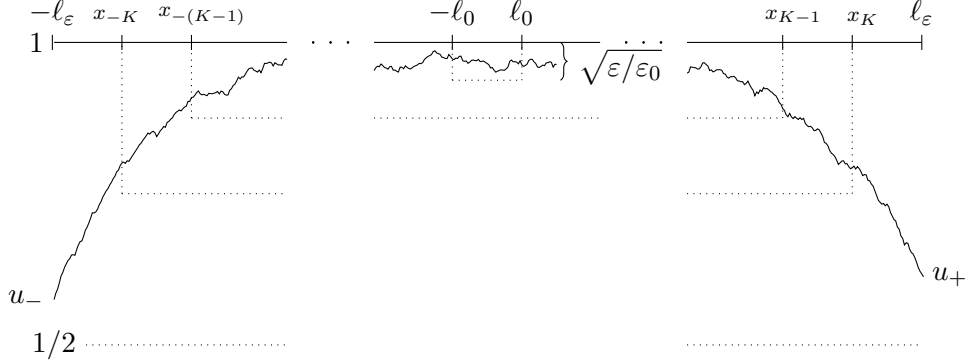


FIGURE 6.2. By iterated rescaling and application of the large deviation bounds we show that the paths relax to a  $O(\varepsilon^{1/2})$  - neighbourhood of 1 within a distance of  $O|\log(\varepsilon)|$ .

On the other hand, on  $\mathcal{A}_0^{\text{bc}}$ , either there exist  $x_- \in [-2\ell, -\ell]$  and  $x_+ \in [\ell, 2\ell]$  such that

$$|u(x_{\pm}) + 1| \leq \delta/2$$

or we have  $u \in [-1 + \delta/2, 1 - \delta/2]$  on an interval of length  $\ell$ . In the latter case, we get easily

$$E_{(-2\ell, 2\ell)}(u) \gtrsim \ell\delta^2.$$

Since this is higher order for  $\ell$  large, we may assume that we are in the former case.

In the former case, we may assume without loss of generality that  $u(x_{\pm}) = -1 + \delta$  and  $u(x_0) = -\delta/2$ . We then use the Modica-Mortola trick to connect the values (a)  $u_-$  and  $u(x_-)$ , (b)  $u(x_-)$  and  $u(x_0)$ , (c)  $u(x_0)$  and  $u(x_+)$ , and (d)  $u(x_+)$  and  $u_+$ . We conclude in the usual way that

$$\inf_{\mathcal{A}_0^{\text{bc}}} E_{(-2\ell, 2\ell)}(u) \geq \varphi_{-1}(u_-) + \varphi_{-1}(u_+) + c_0 - C\delta.$$

Together with (6.52), this completes the proof of Lemma 6.6.  $\square$

## 6.6. Proof of lemmas related to the uniform distribution.

*Proof of Lemma 5.1.* Our argument relies on an iterated rescaling, illustrated in Figure 6.2.

We will define  $K = K_{\varepsilon} \geq 1$  below. We begin by enumerating the partition  $\{x_k\}_{k=-(K+1)}^{K+1}$  of  $(-\ell_{\varepsilon}, \ell_{\varepsilon})$  with width  $2\ell_0$ , so that

$$x_{\pm 1} = \pm\ell_0, x_{\pm 2} = \pm 3\ell_0, \dots, x_{\pm K} = \pm(2K-1)\ell_0, x_{\pm(K+1)} = \pm\ell_{\varepsilon}.$$

For brevity of notation, let

$$\mathcal{A} := \left\{ u : |u(x_k) - 1| \leq \frac{1}{2} \text{ for all } k \in \{-(K+1), -K, \dots, K+1\} \right\}.$$

We will use the elementary facts from probability that for any sets  $A_1$ ,  $A_2$ , and  $A_3$ , we have

$$P(A_1 \cap A_2) \leq P(A_1 \cap A_3) + P(\mathbb{C}A_3 \cap A_2), \quad (6.53)$$

$$P(A_1 \cap A_2 \cap A_3) \leq P(A_1 \cap A_2 | A_3). \quad (6.54)$$

We also use the Markov property from Lemma 3.2 to deduce the following property for conditional measures. If  $\mathcal{A}_{in}$  and  $\tilde{\mathcal{A}}_{in}$  are in  $\mathcal{F}_{[-x_2, x_2]}$  and  $\mathcal{A}_{out}$  is in  $\mathcal{F}_{[-\ell_\varepsilon, -x_2]} \vee \mathcal{F}_{[x_2, \ell_\varepsilon]}$ , then

$$\begin{aligned}
& \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \mathcal{A}_{in} \mid u \in \tilde{\mathcal{A}}_{in} \cap \mathcal{A}_{out} \text{ and } u(\pm x_2) \in (a, b) \right) \\
&= \frac{\mathbb{E}_{(-\ell_\varepsilon, \ell_\varepsilon)}^{\mu_\varepsilon, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}_{out}} \mathbf{1}_{u(\pm x_2) \in (a, b)} \mathbb{E}_{(-x_2, x_2)}^{\mu_\varepsilon, u_-^2, u_+^2} (\mathbf{1}_{\mathcal{A}_{in}} \mathbf{1}_{\tilde{\mathcal{A}}_{in}}) \right)}{\mathbb{E}_{(-\ell_\varepsilon, \ell_\varepsilon)}^{\mu_\varepsilon, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}_{out}} \mathbf{1}_{u(\pm x_2) \in (a, b)} \mathbb{E}_{(-x_2, x_2)}^{\mu_\varepsilon, u_-^2, u_+^2} (\mathbf{1}_{\tilde{\mathcal{A}}_{in}}) \right)} \\
&\leq \sup_{u_\pm^2 \in (a, b)} \mu_{\varepsilon, (-x_2, x_2)}^{u_-^2, u_+^2} (u \in \mathcal{A}_{in} \mid u \in \tilde{\mathcal{A}}_{in}). \tag{6.55}
\end{aligned}$$

Keeping these preliminaries in mind, we now observe that we can make the following decomposition:

$$\begin{aligned}
& \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \mid u \in \mathcal{A} \right) \\
&\stackrel{(6.53)}{\leq} \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \text{ and } \right. \\
&\quad \left. \sup_{x \in [-x_2, x_2]} |u(x) - 1| \leq \frac{1}{2^K} \mid u \in \mathcal{A} \right) \\
&\quad + \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_2, x_2]} |u(x) - 1| \geq \frac{1}{2^K} \mid u \in \mathcal{A} \right).
\end{aligned}$$

For the first term, we can now send the smallness condition into the boundary conditions in the following way:

$$\begin{aligned}
& \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \text{ and } \right. \\
& \quad \left. \sup_{x \in [-x_2, x_2]} |u(x) - 1| \leq \frac{1}{2^K} \middle| u \in \mathcal{A} \right) \\
(6.54) \quad & \leq \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \text{ and } \right. \\
& \quad \left. \sup_{x \in [-x_2, x_2]} |u(x) - 1| \leq \frac{1}{2^K} \middle| \right. \\
& \quad \left. |u(x) - 1| \leq \frac{1}{2^K} \text{ for } x \in \{\pm x_1, \pm x_2\} \text{ and } u \in \mathcal{A} \right) \\
(6.55) \quad & \leq \sup_{u_\pm^2 \in [1-2^{-K}, 1+2^{-K}]} \mu_{\varepsilon, (-x_2, x_2)}^{u_-^2, u_+^2} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \right. \\
& \quad \left. \text{and } \sup_{x \in [-x_2, x_2]} |u(x) - 1| \leq \frac{1}{2^K} \middle| |u(\pm x_1) - 1| \leq \frac{1}{2^K} \right) \\
& \leq \sup_{u_\pm^2 \in [1-2^{-K}, 1+2^{-K}]} \mu_{\varepsilon, (-x_2, x_2)}^{u_-^2, u_+^2} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \middle| \right. \\
& \quad \left. |u(\pm x_1) - 1| \leq \frac{1}{2^K} \right).
\end{aligned}$$

We can iterate this argument to reduce the probability to the form:

$$\begin{aligned}
& \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+} \left( \sup_{x \in [-x_1, x_1]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \middle| u \in \mathcal{A} \right) \\
& \leq \sum_{k=1}^K \sup_{u_\pm^k \in [1-2^{-k}, 1+2^{-k}]} \mu_{\varepsilon, (x_{-(K-k+2)}, x_{K-k+2})}^{u_-^k, u_+^k} \left( \sup_{x \in [x_{-(K-k+1)}, x_{K-k+1}]} |u(x) - 1| \geq \frac{1}{2^{k+1}} \middle| \right. \\
& \quad \left. |u(x_{\pm(K-k+2)}) - 1| \leq \frac{1}{2^k} \right). \tag{6.56}
\end{aligned}$$

Hence it remains to estimate the individual terms in the sum. The argument involves three steps: a large deviation estimate, concatenation, and an iterated rescaling of the deviation of  $u$  from 1.

Step 1: Large deviation estimate. The first step is to derive a uniform large deviation bound for the measures  $\mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{\bar{u}_-, \bar{u}_+}$ . We show that there exists  $C < \infty$  such that for every  $\ell_0 < \infty$  sufficiently large, there exists  $\varepsilon'_0 > 0$  such that for any

$u_{\pm} \in [1/2, 3/2]$  and  $\varepsilon \leq \varepsilon'_0$ , we get

$$\begin{aligned} \mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{u_-, u_+} \left( \sup_{x \in [-\ell_0, \ell_0]} |u(x) - 1| \geq \frac{1}{4} \mid |u(\pm\ell_0) - 1| \leq \frac{1}{2} \right) \\ \leq \exp \left( -\frac{1}{C\varepsilon} \right). \end{aligned} \quad (6.57)$$

In the next steps, we will always assume  $\varepsilon_0 \leq \varepsilon'_0$  to be sufficiently small in this sense, and this is the only restriction on  $\varepsilon_0$  in the proof of the lemma.

To bound the conditional probability in (6.57) it suffices to establish an upper bound on

$$\mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{u_-, u_+} \left( \sup_{x \in [-\ell_0, \ell_0]} |u(x) - 1| \geq \frac{1}{4} \text{ and } |u(\pm\ell_0) - 1| \leq \frac{1}{2} \right) \quad (6.58)$$

and a lower bound on

$$\mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{u_-, u_+} \left( |u(\pm\ell_0) - 1| \leq \frac{1}{2} \right), \quad (6.59)$$

uniformly with respect to  $u_{\pm} \in [1/2, 3/2]$ . To this end, we turn to the uniform large deviation estimates from Propositions 3.4 and 3.5. In fact, we do not even need the second condition in (6.58), and it suffices to bound the probability of the larger set

$$\mathcal{A}_0 := \left\{ u \in C([-3\ell_0, 3\ell_0]) : u(\pm 3\ell_0) = u_{\pm}, \sup_{x \in [-\ell_0, \ell_0]} |u(x) - 1| \geq \frac{1}{4} \right\},$$

The estimate (3.17) gives that for any  $\gamma, \delta > 0$ , we have for sufficiently small  $\varepsilon$  that

$$\mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{u_-, u_+}(\mathcal{A}_0) \leq \exp \left( -\frac{1}{\varepsilon} (\Delta E(B(\mathcal{A}_0, \delta)) - \gamma) \right), \quad (6.60)$$

where  $\Delta E$  is defined in (3.13) and

$$\mathcal{A}^{\text{bc}} = \{ u \in C([-3\ell_0, 3\ell_0]) : u(\pm 3\ell_0) = u_{\pm} \}.$$

Consider now a small  $\delta > 0$  to be fixed below and a function  $u \in B(\mathcal{A}_0, \delta)$ . Because the boundary conditions are in  $[1/2, 3/2]$  and  $\ell_0$  is large, the infimum of the energy must take place over functions such that

$$\max \left\{ \min_{x \in [-3\ell_0, -\ell_0]} |u(x) - 1|, \min_{x \in [\ell_0, 3\ell_0]} |u(x) - 1| \right\} \lesssim \frac{1}{\sqrt{\ell_0}}. \quad (6.61)$$

(Indeed,  $u$  must be close to either 1 or  $-1$  at some point in each of the intervals, and if  $u$  were instead close to  $-1$  on either interval, satisfying the boundary conditions would lead to an even greater energetic cost than the one we will arrive at below.) Let us label the minimizing points  $x_-$  and  $x_+$ . Moreover, let us define  $x_*$  to be a point in  $(-\ell_0, \ell_0)$  such that

$$|u(x_*) - 1| \geq \frac{1}{4} - \delta.$$

As above in Subsection 6.1, we now define  $\varphi(u) := |\int_u^1 \sqrt{2V(s)} ds|$  and apply the “Modica-Mortola trick” on  $(-3\ell_0, x_-)$ ,  $(x_-, x_*)$ ,  $(x_*, x_+)$ , and  $(x_+, 3\ell_0)$  to

recover

$$\begin{aligned}
\inf_{B(\mathcal{A}_0, \delta)} E(u) &\geq \varphi(u_-) - \varphi(u(x_-)) + \varphi(u(x_*)) - \varphi(u(x_-)) \\
&\quad + \varphi(u(x_*)) - \varphi(u(x_+)) + \varphi(u_+) - \varphi(u(x_+)) \\
&\stackrel{(6.61)}{\geq} 2\varphi(u(x_*)) + \varphi(u_-) + \varphi(u_+) - o(1)_{\ell_0 \uparrow \infty} \\
&\geq 2\varphi_{1/4} + \varphi(u_-) + \varphi(u_+) - o(1)_{\ell_0 \uparrow \infty} - o(1)_{\delta \downarrow 0}, \quad (6.62)
\end{aligned}$$

where

$$\varphi_{1/4} := \min\{\varphi(3/4), \varphi(5/4)\}.$$

On the other hand, a standard construction gives

$$\inf_{\mathcal{A}^{\text{bc}}} E(u) \leq \varphi(u_-) + \varphi(u_+) + o(1)_{\ell_0 \uparrow \infty}. \quad (6.63)$$

Now fixing  $\delta > 0$  and  $\gamma > 0$  sufficiently small, the combination of (6.60), (6.62), and (6.63) gives for sufficiently small  $\varepsilon$  that

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(\mathcal{A}_0) \leq \exp\left(-\frac{3/2 \varphi_{1/4}}{\varepsilon}\right). \quad (6.64)$$

We now remark that the lower bound on (6.59) follows easily from Proposition 3.5. Indeed, for a fixed  $0 < \delta < \frac{1}{2}$ , the set of interest can be written as the  $\delta$  ball around the set  $\mathcal{A}_1$  defined as

$$\mathcal{A}_1 := \left\{ u : u(\pm 3\ell_0) = u_{\pm}, \quad |u(\pm \ell_0) - 1| \leq \frac{1}{2} - \delta \right\}.$$

We recover for any  $\gamma > 0$  and for  $\varepsilon > 0$  sufficiently small that

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(B(\mathcal{A}_1, \delta)) \geq \exp\left(-\frac{1}{\varepsilon}(\Delta E(\mathcal{A}_1) + \gamma)\right), \quad (6.65)$$

where  $\Delta E$  is defined in (3.13). The constraint in  $\mathcal{A}_1$  is inactive/slack in the optimization for  $\ell_0$  sufficiently large, and the usual construction together with the usual Modica-Mortola estimate thus gives

$$\Delta E(\mathcal{A}_1) \leq o(1)_{\ell_0 \uparrow \infty}.$$

Plugging back into (6.65) gives

$$\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}(B(\mathcal{A}_1, \delta)) \geq \exp\left(-\frac{2\gamma}{\varepsilon}\right),$$

which together with (6.64) gives (6.57) with  $C = 1/\varphi_{1/4}$  as long as  $\gamma$  is chosen sufficiently small.

Step 2: Concatenation. The next step is to prove for any  $K \in \mathbb{N}$  that

$$\begin{aligned}
\mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+} \left( \sup_{x \in [x_{-K}, x_K]} |u(x) - 1| \geq \frac{1}{4} \mid u \in \mathcal{A} \right) \\
\leq 2K \exp\left(-\frac{1}{C\varepsilon}\right), \quad (6.66)
\end{aligned}$$

uniformly for  $u_{\pm} \in (1/2, 3/2)$ . As usual, the idea is to break up the larger interval by conditioning on the boundary values. The restriction of the boundary values on each subinterval to  $(1/2, 3/2)$  because of  $u \in \mathcal{A}$  will allow us to apply the uniform estimate from Step 1.

We will consider the non-overlapping subintervals  $[x_k, x_{k+1}]$  for  $k = -K, \dots, K-1$ .  
1. Decomposing the interval in this way gives

$$\begin{aligned} & \mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+} \left( \sup_{x \in [x_{-K}, x_K]} |u(x) - 1| \geq \frac{1}{4} \mid u \in \mathcal{A} \right) \\ & \leq \sum_{k=1}^{2K} \mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+} \left( \sup_{x \in [x_k, x_{k+1}]} |u(x) - 1| \geq \frac{1}{4} \mid u \in \mathcal{A} \right). \end{aligned} \quad (6.67)$$

Now the Markov property implies that for  $k \in \{-K, \dots, K-1\}$  we have

$$\begin{aligned} & \mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+} \left( \sup_{x \in [x_k, x_{k+1}]} |u(x) - 1| \geq \frac{1}{4} \text{ and } u \in \mathcal{A} \right) \\ & = \mathbb{E}_{(x_{-(K+1)}, x_{K+1})}^{\mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+}} \left( \prod_{j=-K}^K \mathbf{1}_{\{|u(x_j) - 1| \leq \frac{1}{2}\}} \mathbf{1}_{\{\sup_{x \in [x_k, x_{k+1}]} |u(x) - 1| \geq \frac{1}{4}\}} \right) \\ & \stackrel{(6.55)}{\leq} \sup_{u_{\pm}^k \in [1/2, 3/2]} \mu_{\varepsilon, (x_{k-1}, x_{k+2})}^{u_-, u_+^k} \left( \sup_{x \in [x_k, x_{k+1}]} |u(x) - 1| \geq \frac{1}{4} \right. \\ & \quad \left. \mid \max\{|u(x_k) - 1|, |u(x_{k+1}) - 1|\} \leq \frac{1}{2} \right) \mu_{\varepsilon, (x_{-(K+1)}, x_{K+1})}^{u_-, u_+}(\mathcal{A}). \end{aligned}$$

Hence, using the translational invariance of the measures  $\mu_{\varepsilon, (x_-, x_+)}^{u_-, u_+}$ , we bound the right-hand side of equation (6.67) by

$$\begin{aligned} & 2K \sup_{u_{\pm}^k \in [1/2, 3/2]} \mu_{\varepsilon, (-3\ell_0, 3\ell_0)}^{u_-, u_+} \left( \sup_{x \in [-\ell_0, \ell_0]} |u(x) - 1| \geq \frac{1}{4} \mid |u(\pm\ell_0) - 1| \leq \frac{1}{2} \right) \\ & \stackrel{(6.57)}{\leq} 2K \exp\left(-\frac{1}{C\varepsilon}\right), \end{aligned}$$

which is what we wanted to show.

**Step 3: Rescaling and iteration.** In this step, we rescale the deviation of  $u$  from 1. We fix  $k$  and consider the random variables

$$\hat{u} := 2^{k-1}(u - 1) + 1.$$

For  $u$  distributed according to  $\mu_{\varepsilon, (-\ell, \ell)}^{u_-, u_+}$ , the profile  $\hat{u}$  is distributed according to a rescaled version of the measure. Indeed, the Radon-Nikodym density with respect to the Brownian bridge measure with *modified noise strength*  $\hat{\varepsilon} := 4^{k-1}\varepsilon$  and rescaled boundary conditions is proportional to

$$\exp\left(-\frac{1}{\hat{\varepsilon}} \int_{-\ell}^{\ell} 4^{k-1} V\left(2^{-(k-1)}(\hat{u} - 1) + 1\right) dx\right). \quad (6.68)$$

Let us give a name to the modified potential

$$\hat{V}(\hat{u}) := 4^{k-1} V(2^{-(k-1)}(\hat{u} - 1) + 1)$$

and the associated energy

$$\hat{E}(\hat{u}) := \int_{-\ell}^{\ell} \left( \frac{1}{2} (\partial_x \hat{u})^2 + \hat{V}(\hat{u}) \right) dx.$$

We now make a series of observations that will allow us to apply the same large deviation bounds from Steps 1 and 2 to the rescaled random variables  $\hat{u}$ .

First consider how the sets involved in (6.56) behave under the rescaling. Notice that  $u$  satisfies

$$\sup_{x \in [x_{-(K-k+2)}, x_{K-k+2}]} |u(x) - 1| \geq \frac{1}{2^{k+1}}$$

precisely when

$$\sup_{x \in [x_{-(K-k+2)}, x_{K-k+2}]} |\hat{u}(x) - 1| \geq \frac{1}{4}.$$

Similarly, for the set on which we condition, we have that  $u$  satisfies

$$|u(x_j) - 1| \leq \frac{1}{2^k} \text{ for } j \in \{-(K-k+3), \dots, K-k+3\}$$

precisely when  $\hat{u}$  satisfies

$$|\hat{u}(x_j) - 1| \leq \frac{1}{2} \text{ for } j \in \{-(K-k+3), \dots, K-k+3\}.$$

Hence each term in (6.56) can be bounded if we can establish that the bound from Step 2 also holds for the measure governing  $\hat{u}$ .

In order to show that the estimates from Step 1 and 2 hold uniformly for the measure of the rescaled random variables  $\hat{u}$ , we need to be able to invoke Propositions 3.4 and 3.5 (with uniform constants). This in turn requires uniform control on the boundary values, the minimum energy  $\hat{E}$  over the sets of interest, and the Lipschitz constant of  $\hat{V}$ . The boundary values are easy: On the sets of interest, the boundary values  $u_{\pm} \in (1/2, 3/2)$ . On the other hand, the minimum of the energy  $\hat{E}$  is bounded uniformly with respect to  $k$  on the sets of interest. Indeed, consider

$$\mathcal{C} := \left\{ u : |u(x_j) - 1| \leq \frac{1}{2^k} \text{ for } j \in \{-(K-k+3), \dots, K-k+3\} \right. \\ \left. \text{and } \sup_{x \in [x_{-(K-k+2)}, x_{K-k+2}]} |u(x) - 1| \geq \frac{1}{2^{k+1}} \right\}$$

and let  $\hat{\mathcal{C}}$  denote the image of the set under the transformation  $u \rightarrow \hat{u}$ . By the usual method (“Modical Mortola trick” for the lower bound and construction for the upper bound), one can check that there exists  $R < \infty$  such that, for every  $k \in \mathbb{N}$ , one has

$$\inf_{\hat{u} \in \hat{\mathcal{C}}} \hat{E}(\hat{u}) = 4^{k-1} \inf_{u \in \mathcal{C}} E(u) \leq R.$$

Finally, because of Assumption 1.1, we have a uniform bound on the Lipschitz constant of  $V$ . Indeed, let  $C := 3/2 + 2\ell_0 R + 1$ . Then uniformly with respect to  $k \in \mathbb{N}$ , the potential  $\hat{V}$  satisfies

$$\sup_{|\hat{u}| \leq C} |\hat{V}'(\hat{u})| \leq \sup_{|u-1| \leq 2^{-k+1}(C+1)} 2^{k-1} |V'(u)| \\ \leq \sup |V''(\tau)|(C+1),$$

where the supremum is taken over  $\tau \in [1 - 2(C+1), 1 + 2(C+1)]$ .

Hence, the potential satisfies the requirements of Propositions 3.4 and 3.5. The remaining requirement in order to invoke large deviation theory is that

$$4^{k-1} \varepsilon \leq \varepsilon_0 \quad \text{for all } k \leq K,$$

which is true if

$$4^{K-1}\varepsilon \leq \varepsilon_0.$$

Therefore we choose  $K$  to be an integer satisfying

$$\frac{1}{2^{K+1}} \leq \sqrt{\frac{\varepsilon}{\varepsilon_0}} \leq \frac{1}{2^{K-1}}. \quad (6.69)$$

With the restriction (6.69) on  $K$ , the arguments used in Step 1 and Step 2 carry over to the rescaled measures governing the  $\hat{u}$ .

We are now ready to complete the argument. Indeed, recalling the decomposition from (6.56), we have

$$\begin{aligned} & \mu_{\varepsilon,(-\ell_\varepsilon,\ell_\varepsilon)}^{u_-,u_+} \left( \sup_{x \in [-\ell_0,\ell_0]} |u(x) - 1| \geq \sqrt{\frac{\varepsilon}{\varepsilon_0}} \mid u \in \mathcal{A} \right) \\ (6.69) \quad & \leq \mu_{\varepsilon,(-\ell_\varepsilon,\ell_\varepsilon)}^{u_-,u_+} \left( \sup_{x \in [-\ell_0,\ell_0]} |u(x) - 1| \geq \frac{1}{2^{K+1}} \mid u \in \mathcal{A} \right) \\ (6.56) \quad & \leq \sum_{k=1}^K \sup_{u_\pm^k \in [1-2^{-k}, 1+2^{-k}]} \\ & \mu_{\varepsilon,(x_{-(K-k+2)}, x_{K-k+2})}^{u_-^k, u_+^k} \left( \sup_{x \in [x_{-(K-k+1)}, x_{K-k+1}]} |u(x) - 1| \geq \frac{1}{2^{k+1}} \mid \right. \\ & \left. |u(x_{\pm(K-k+2)}) - 1| \leq \frac{1}{2^k} \right). \end{aligned} \quad (6.70)$$

From the preceding argument, we can now apply the estimate (6.66) for the rescaled measures to bound the  $k^{th}$  summand above by

$$2(K-k+1) \exp \left( -\frac{1}{C 4^{k-1} \varepsilon} \right).$$

Substituting into the right-hand side of (6.70), we deduce

$$\begin{aligned} & \mu_{\varepsilon,(-\ell_\varepsilon,\ell_\varepsilon)}^{u_-,u_+} \left( \sup_{x \in [-\ell_0,\ell_0]} |u(x) - 1| \geq \sqrt{\frac{\varepsilon}{\varepsilon_0}} \mid u \in \mathcal{A} \right) \\ & \leq \sum_{k=1}^K 2(K-k+1) \exp \left( -\frac{1}{C 4^{k-1} \varepsilon} \right) \\ & = 2 \sum_{k=0}^{K-1} (K-k) \left( \exp \left( -\frac{1}{C 4^k \varepsilon} \right) \right)^{4^{K-k}} \\ (6.69) \quad & \leq 2 \sum_{k=0}^{K-1} (K-k) r^{4^{K-k}}, \quad \text{for } r := \exp \left( -\frac{1}{C \varepsilon_0} \right) \\ & \leq 2 \sum_{k'=1}^{\infty} k' r^{k'} = \frac{2r}{(r-1)^2} \leq 4r \quad \text{for } r \in (0, 1/4]. \end{aligned}$$

□



*Proof of Lemma 5.2.* We start by defining some sets. We denote the set of paths that we condition on by

$$\mathcal{A} := \{u \in C([- \ell_\varepsilon, \ell_\varepsilon]) : |u(\pm(2k-1)\ell_0) - 1| \leq \frac{1}{2}, k = 1, 2, \dots, K_\varepsilon\}.$$

For  $\varepsilon, \varepsilon_0 > 0$  let us also fix the following subset of  $\mathcal{A}$

$$\mathcal{A}_\varepsilon := \left\{ u \in \mathcal{A} : |u(\pm\ell_0) - 1| \leq \left( \frac{\varepsilon}{\varepsilon_0} \right)^{1/2} \right\}.$$

Then Lemma 5.1 implies in particular that, for a small but fixed  $\varepsilon_0 > 0$  and for  $\varepsilon \leq \varepsilon_0$ , we have

$$\mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+}(\mathcal{A}_\varepsilon) \geq \frac{1}{2} \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+}(\mathcal{A}). \quad (6.71)$$

From now on, we fix an  $\varepsilon_0$  such that this identity holds. This will be the only restriction on  $\varepsilon_0$ .

Let us also introduce a notation for the set of paths that have a hitting point of 1 in  $[-\ell_0, \ell_0]$

$$\mathcal{B} := \{u \in C([- \ell_\varepsilon, \ell_\varepsilon]) : \exists x \in [-\ell_0, \ell_0] \text{ such that } u(x) = 1\}.$$

As a slight abuse of notation we will use the same letter  $\mathcal{B}$  to denote the set of paths  $u \in \mathcal{B}$  restricted to  $[-\ell_0, \ell_0]$ .

Using the Markov property (3.7), we get for any  $u^\pm \in [1/2, 3/2]$  that

$$\begin{aligned} \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+}(\mathcal{A} \cap \mathcal{B}) &\geq \mu_{\varepsilon, (-\ell_\varepsilon, \ell_\varepsilon)}^{u_-, u_+}(\mathcal{A}_\varepsilon \cap \mathcal{B}) \\ &= \mathbb{E}_{(-\ell_\varepsilon, \ell_\varepsilon)}^{\mu_\varepsilon, u_-, u_+} \left( \mathbf{1}_{\mathcal{A}_\varepsilon}(u) \mu_{\varepsilon, (-\ell_0, \ell_0)}^{u(-\ell_0), u(+\ell_0)}(\mathcal{B}) \right). \end{aligned} \quad (6.72)$$

Our main task is thus to derive a lower bound for the probabilities

$$\mu_{\varepsilon, (-\ell_0, \ell_0)}^{u_-, u_+}(\mathcal{B}) \quad (6.73)$$

that holds uniformly in the boundary conditions. In view of the definition of  $\mathcal{A}_\varepsilon$ , it is sufficient to consider boundary conditions  $u_\pm$  that are  $O(\varepsilon^{1/2})$  close to 1:

$$1 - \left( \frac{\varepsilon}{\varepsilon_0} \right)^{1/2} \leq u_\pm \leq 1 + \left( \frac{\varepsilon}{\varepsilon_0} \right)^{1/2}. \quad (6.74)$$

As in the proof of Lemma 5.1, we rescale the random profile  $u$  around 1, this time by a factor  $\varepsilon^{-1/2}$ . More precisely, we consider the transformation

$$\hat{u}(x) := \varepsilon^{-1/2}(u(x) - 1) + 1.$$

According to its definition, a path  $u$  is in the set  $\mathcal{B}$  if and only if  $\hat{u}$  is in  $\mathcal{B}$ . Hence, we can express the probability (6.73) in terms of  $\hat{u}$ .

The random variable  $\hat{u}$  is distributed according to a rescaled version of  $\mu_{\varepsilon, (-\ell_0, \ell_0)}^{u_-, u_+}$ . The variance of the Gaussian reference measure becomes one and the rescaled boundary values are

$$\hat{u}_\pm := \varepsilon^{-1/2}(u_\pm - 1) + 1.$$

Note that the condition (6.74) implies that these rescaled boundary conditions take values in an order-one interval around 1. More precisely, the distribution of  $\hat{u}$  is absolutely continuous with respect to  $\mathcal{W}_{1, (-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}$  and the Radon Nikodym density

of the rescaled measure is proportional to  $\exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right)$ , where  $\hat{V}(\hat{u}) := \frac{1}{\varepsilon} V(\varepsilon^{1/2}(\hat{u} - 1) + 1)$ . Hence we can rewrite

$$\begin{aligned} & \mu_{\varepsilon, (-\ell_0, \ell_0)}^{u_-, u_+}(\mathcal{B}) \\ &= \frac{\mathbb{E}_{(-\ell_0, \ell_0)}^{\mathcal{W}_1, \hat{u}_-, \hat{u}_+} \left( \mathbf{1}_{\mathcal{B}}(\hat{u}) \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right) \right)}{\mathbb{E}_{(-\ell_0, \ell_0)}^{\mathcal{W}_1, \hat{u}_-, \hat{u}_+} \left( \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right) \right)}. \end{aligned} \quad (6.75)$$

The denominator of this expression can be trivially bounded above by 1. To get a lower bound for the numerator, we can write for example

$$\begin{aligned} & \mathbb{E}_{(-\ell_0, \ell_0)}^{\mathcal{W}_1, \hat{u}_-, \hat{u}_+} \left( \mathbf{1}_{\mathcal{B}}(\hat{u}) \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right) \right) \\ & \geq \mathbb{E}_{(-\ell_0, \ell_0)}^{\mathcal{W}_1, \hat{u}_-, \hat{u}_+} \left( \mathbf{1}_{\mathcal{B}^*}(\hat{u}) \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right) \right) \\ & \geq \mathcal{W}_{1, (-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}(\mathcal{B}^*) \inf_{\hat{u} \in \mathcal{B}^*} \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right). \end{aligned} \quad (6.76)$$

Here we have made the probability smaller by restricting the integration to the set

$$\mathcal{B}^* := \left\{ \hat{u} \in \mathcal{B} : \sup_{x \in [-\ell_0, \ell_0]} |\hat{u}(x) - 1| \leq 3\varepsilon_0^{-1/2} \right\}.$$

Using the translation invariance of the Gaussian measures, we can get a lower bound on the Gaussian probabilities that holds uniformly in the boundary conditions. For example, set

$$\begin{aligned} \mathcal{B}^{**} &:= \left\{ \hat{u} \in C([-\ell_0, \ell_0]) : \sup_{x \in [-\ell_0, \ell_0]} \hat{u}(x) \in (\varepsilon_0^{-1/2}, 2\varepsilon_0^{-1/2}), \right. \\ & \quad \left. \text{and } \inf_{x \in [-\ell_0, \ell_0]} \hat{u}(x) \in (-2\varepsilon_0^{-1/2}, -\varepsilon_0^{-1/2}) \right\}. \end{aligned}$$

Then, on the one hand, for every path  $\hat{u} \in \mathcal{B}^{**}$  and for all  $u_{\varepsilon, \pm} \in [1 - \varepsilon_0^{-1/2}, 1 + \varepsilon_0^{-1/2}]$ , the shifted paths  $\hat{u} + h_{(-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}$  lies in  $\mathcal{B}^*$ . (Recall the definition (3.4) of the affine profile  $h_{(-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}$ ). On the other hand, by definition, shifting by  $h_{(-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}$  transforms the measure  $\mathcal{W}_{1, (-\ell_0, \ell_0)}^{0,0}$  into  $\mathcal{W}_{1, (-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}$ . This implies that

$$\mathcal{W}_{1, (-\ell_0, \ell_0)}^{\hat{u}_-, \hat{u}_+}(\mathcal{B}^*) \geq \mathcal{W}_{1, (-\ell_0, \ell_0)}^{0,0}(\mathcal{B}^{**}) =: c > 0. \quad (6.77)$$

Hence it remains to get a lower bound on the second term in (6.76). As above in the proof of Lemma 5.1, Assumption 1.1 on  $V$  and Taylor's formula imply that  $\hat{V}$  satisfies

$$\sup_{|u-1| \leq 3\varepsilon_0^{-1/2}} \sup_{\varepsilon \in (0,1)} \varepsilon^{-1} V(\varepsilon^{1/2}(u-1) + 1) =: C < \infty.$$

Plugging this into (6.76), we get

$$\inf_{\hat{u} \in \mathcal{B}^*} \exp\left(-\int_{-\ell_0}^{\ell_0} \hat{V}(\hat{u}) dx\right) \geq \exp(-2C\ell_0).$$

Hence, summarizing this calculation, we get uniformly for all  $u_{\pm}$  satisfying (6.74) that

$$\mu_{\varepsilon,(-\ell_0,\ell_0)}^{u_-,u_+}(\mathcal{B}) \geq c \exp(-2C\ell_0).$$

Finally, plugging this back into (6.72), we get

$$\begin{aligned} \mu_{\varepsilon,(-\ell_{\varepsilon},\ell_{\varepsilon})}^{u_-,u_+}(\mathcal{A} \cap \mathcal{B}) &\geq c \exp(-2C\ell_0) \mu_{\varepsilon,(-\ell_{\varepsilon},\ell_{\varepsilon})}^{u_-,u_+}(\mathcal{A}_{\varepsilon}) \\ &\stackrel{(6.71)}{\geq} \frac{1}{2} c \exp(-2C\ell_0) \mu_{\varepsilon,(-\ell_{\varepsilon},\ell_{\varepsilon})}^{u_-,u_+}(\mathcal{A}). \end{aligned}$$

Thus we get the desired conclusion for  $1 - \lambda := \frac{1}{2} c \exp(-2C\ell_0)$ .  $\square$

*Proof of Lemma 5.4. Step 1.* We begin by ruling out long layers to the left and to the right of  $Y$ . Once we know that layers are bounded in length, we can use a reflection argument as in the proof of Theorem 1.5 to turn them into wasted excursions and estimate their probability. To this end, we define the set  $\mathcal{A}_{Y,2}$  of functions that are bounded away from  $\pm 1$  on a whole subinterval outside of  $Y$ :

$$\begin{aligned} \mathcal{A}_{Y,2} := \{u \in J_Y : &\text{ there exists a } k \text{ with} \\ &k \leq k_- \text{ or } k \geq k_+ - 1 \text{ such that} \\ &u \in [-1 + \delta, 1 - \delta] \text{ on all of } [x_k, x_{k+1}]\}. \end{aligned}$$

As usual, we note that  $\mathcal{A}_{Y,3}$  is contained within  $\mathcal{A}_{Y,2} \cup (\mathcal{A}_{Y,3} \cap \mathcal{A}_{Y,2})$ . Our first step is to show that (5.6) holds for  $\mathcal{A}_{Y,2}$ . In fact,  $\mathcal{A}_{Y,2}$  is of higher order for  $M$  and  $\delta^2 \ell$  sufficiently large.

The set  $\mathcal{A}_{Y,2}$  can be written in the obvious way as the union of sets  $\mathcal{A}_{Y,2}^k$  that have bad behavior on a given subinterval  $[x_k, x_{k+1}]$ . Without loss of generality, suppose that  $k \leq k_-$ .

Then we introduce the following sets for a Markovian decomposition:

$$\begin{aligned} \mathcal{A}_k^{\ominus} &:= \{u : |u(x_j)| \leq M \text{ for all } j \leq k-1\}, \\ \mathcal{A}_k^{\oplus} &:= \{u : |u(x_j)| \leq M \text{ for all } j \geq k+2, \\ &\text{and at least one } \delta^- \text{ up layer } \leq 2\ell \text{ in } Y\}, \\ \mathcal{A}_k^{\odot} &:= \{u : |u(x_j)| \leq M \text{ for } j = k-1, \dots, k+2\}, \\ \mathcal{A}_{\delta,k}^{\odot} &:= \{u \in \mathcal{A}_k^{\odot} : u \in [-1 + \delta, 1 - \delta] \text{ on all of } [x_k, x_{k+1}]\}. \end{aligned}$$

We remark that

$$\begin{aligned} \mathcal{A}^{\ominus} &\in \mathcal{F}_{[-L_{\varepsilon}, x_{k-1}]}, \quad \text{and} \quad \mathcal{A}^{\oplus} \in \mathcal{F}_{[x_{k+2}, L_{\varepsilon}]}, \\ \text{while } \mathcal{A}^{\odot} &\in \mathcal{F}_{[x_{k-1}, x_{k+2}]} \quad \text{and} \quad \mathcal{A}_{\delta,k}^{\odot} \in \mathcal{F}_{[x_{k-1}, x_{k+2}]}. \end{aligned}$$

Consequently, the decompositions  $\mathcal{A}_{Y,2}^k = \mathcal{A}_k^{\ominus} \cap \mathcal{A}_{\delta,k}^{\odot} \cap \mathcal{A}_k^{\oplus}$  and  $\mathcal{J}_Y = \mathcal{A}_k^{\ominus} \cap \mathcal{A}_k^{\odot} \cap \mathcal{A}_k^{\oplus}$  lend themselves to an application of the Markov property from Lemma 3.2. We will often use such decompositions in the proofs below.

In the proof at hand, the Markov property from Lemma 3.2 gives

$$\mu_{\varepsilon,(-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}(\mathcal{A}_{Y,2}^k) \leq \sup_{u_{\pm} \in [-M, M]} \frac{\mathbb{E}_{(x_{k-1}, x_{k+2})}^{\mu_{\varepsilon}, u_-, u_+}(\mathbf{1}_{\mathcal{A}_{\delta,k}^{\odot}})}{\mathbb{E}_{(x_{k-1}, x_{k+2})}^{\mu_{\varepsilon}, u_-, u_+}(\mathbf{1}_{\mathcal{A}_k^{\odot}})} \mu_{\varepsilon,(-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}(\mathcal{J}_Y).$$

It suffices to bound the ratio of expectations on the right-hand side. For the denominator, we observe that

$$\inf_{u_{\pm} \in [-M, M]} \mathbb{E}_{(x_{k-1}, x_{k+2})}^{\mu_{\varepsilon}, u_{-}, u_{+}}(\mathbf{1}_{\mathcal{A}_k^{\odot}}) \gtrsim 1 \quad (6.78)$$

for  $M$  sufficiently large. In fact, this bound follows immediately from the large deviation bound (3.17) and a simple energy estimate applied to the complement.

Hence, it suffices to bound the numerator. Recalling the bound (4.12), the expectation in the numerator can be estimated by

$$\exp\left(-\frac{1}{\varepsilon}\left(\frac{\delta^2 \ell}{C_1} - 2\gamma\right)\right) \leq \exp\left(-\frac{\delta^2 \ell - 1}{\varepsilon C_1}\right).$$

For  $\delta^2 \ell$  sufficiently large, this drops below the threshold expressed in the exponential in (5.7). Hence, summing the probabilities of  $\mathcal{A}_{Y,2}^k$  over  $k$ , the probability of  $\mathcal{A}_{Y,2}$  is negligible in the sense that, in order to establish (5.7), it suffices to show that it holds for  $\tilde{\mathcal{A}}_{Y,3} := \mathcal{A}_{Y,3} \setminus \mathcal{A}_{Y,2}$ . For ease of notation, we drop the tildes for the remainder of the proof of the lemma.

Step 2. We will now show the desired bound for  $\mathcal{A}_{Y,3}$ . That is, we will show that for any  $\gamma > 0$  there exists an  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  we have

$$\mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}(\mathcal{A}_{Y,3}) \lesssim L_{\varepsilon} \exp\left(-\frac{c_0 - \gamma}{\varepsilon}\right) \mu_{\varepsilon, (-L_{\varepsilon}, L_{\varepsilon})}^{-1,1}(\mathcal{J}_Y).$$

The proof uses a reflection argument very similar to the argument in the proof of the upper bound in Theorem 1.5.

As above in (5.5) the set  $\mathcal{A}_{Y,3}$  can be expressed as  $\mathcal{A}_{Y,3} = \mathcal{A}_{Y,3}^{-} \cup \mathcal{A}_{Y,3}^{+}$  where

$$\begin{aligned} \mathcal{A}_{Y,3}^{-} &= \{u \in \mathbb{C}\mathcal{A}_1 \cap \mathbb{C}\mathcal{A}_{Y,2} : u \text{ has a } \delta^{-} \text{ up layer} \\ &\quad \text{contained in } [-L_{\varepsilon}, k_{-}\ell] \text{ and a } \delta^{-} \text{ up layer } \leq 2\ell \text{ in } J_Y\}, \\ \mathcal{A}_{Y,3}^{+} &= \{u \in \mathbb{C}\mathcal{A}_1 \cap \mathbb{C}\mathcal{A}_{Y,2} : u \text{ has a } \delta^{-} \text{ down layer} \\ &\quad \text{contained in } [k_{+}\ell, L_{\varepsilon}] \text{ and a } \delta^{-} \text{ up layer } \leq 2\ell \text{ in } J_Y\}. \end{aligned}$$

We will only give the bound for the set  $\mathcal{A}_{Y,3}^{-}$ . The proof of the corresponding bound for  $\mathcal{A}_{Y,3}^{+}$  follows in the same way. The set  $\mathcal{A}_{Y,3}^{-}$  is contained in the union of  $k$  from  $-(N_{\varepsilon} - 1)$  to  $k_{-}$  of the sets

$$\begin{aligned} \mathcal{A}_{Y,3}^{-,k} &:= \{u \in \mathbb{C}\mathcal{A}_1 : u \text{ has a } \delta^{-} \text{ up layer} \\ &\quad \text{contained in } [x_{k-1}, x_{k+1}] \text{ and a } \delta^{-} \text{ up layer } \leq 2\ell \text{ in } Y\}. \end{aligned}$$

As in the proof of Theorem 1.5, we will transform the additional  $\delta^{-}$  transition layer into a wasted  $\delta^{-}$  excursion to control the probability. We need to reflect in such a way as to (a) create a wasted excursion in  $[x_{k-1}, x_{k+1}]$  and (b) leave at least one  $\delta^{-}$  up layer in  $Y$ . To this end, we define the left stopping point capturing the additional  $\delta^{-}$  up layer

$$\begin{aligned} \chi_{-} &= \inf\{x > x_{k-1} : u(x) = 0 \\ &\quad \text{and } u(y_1) = -1 + \delta \text{ for some } y_1 \in (x_{k-1}, x)\} \end{aligned}$$

and the right stopping point

$$\begin{aligned} \chi_{+} &:= \sup\{x \leq y_{+} : u(x) = 0 \text{ and there exist } y_1 < y_2 \\ &\quad \text{both in } (x, y_{+}) \text{ with } u(y_1) = -1 + \delta, u(y_2) = 1 - \delta\}, \end{aligned}$$

where  $y_+ := \sup Y$  is the right boundary of  $Y$ . As before we will use the convention that  $\chi_\pm = \mp L_\varepsilon$  if these sets are empty. As in the proof of Theorem 1.5, the reflection operator

$$R = R_{\chi_-}^{\chi_+},$$

reflects the paths  $u$  between the stopping points  $\chi_\pm$  while preserving  $\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}$ . On the other hand, it maps the set  $\mathcal{A}_{Y,3}^{-,k}$  into the set

$$\begin{aligned} \hat{\mathcal{A}}_{Y,3}^{-,k} &:= \{u \in \mathbb{C}\mathcal{A}_1 : \text{at least one wasted } \delta^- \text{ excursion in } [x_{k-1}, x_{k+1}] \\ &\quad \text{and at least one } \delta^- \text{ up layer } \leq 2\ell \text{ in } \mathcal{J}_Y\}. \end{aligned}$$

Hence, the estimate (5.7) will follow if we can establish, uniformly in  $k$ , that

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( \hat{\mathcal{A}}_{Y,3}^{-,k} \right) \leq \exp \left( -\frac{c_0 - \gamma}{\varepsilon} \right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_Y), \quad (6.79)$$

which will follow from the Markov property and a large deviation estimate. Indeed, let us define the following sets:

$$\begin{aligned} \mathcal{A}_k^\ominus &:= \{u : |u(x_j)| \leq M \text{ for all } j \leq k-2\}, \\ \mathcal{A}_k^\oplus &:= \{u : |u(x_j)| \leq M \text{ for all } j \geq k+2 \text{ and at least one } \delta^- \text{ up layer } \leq 2\ell \text{ in } Y\}, \\ \mathcal{A}_k^\odot &:= \{u : |u(x_j)| \leq M \text{ for } j = k-2, \dots, k+2\}, \\ \mathcal{A}_{w,k}^\odot &:= \{u \in \mathcal{A}_k^\odot : u \text{ has a wasted } \delta^- \text{ excursion in } [x_{k-1}, x_{k+1}]\}. \end{aligned}$$

Then we can decompose  $\hat{\mathcal{A}}_{Y,3}^{-,k} = \mathcal{A}_k^\ominus \cap \mathcal{A}_{w,k}^\odot \cap \mathcal{A}_k^\oplus$  and  $\mathcal{J}_Y = \mathcal{A}_k^\ominus \cap \mathcal{A}_k^\odot \cap \mathcal{A}_k^\oplus$ , so that applying the Markov property as in Lemma 3.2 gives

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( \hat{\mathcal{A}}_{Y,3}^{-,k} \right) \leq \sup_{u_\pm \in [-M,M]} \frac{\mathbb{E}_{(x_{k-2}, x_{k+2})}^{\mu_\varepsilon, u_-, u_+} (\mathbf{1}_{\mathcal{A}_{w,k}^\odot})}{\mathbb{E}_{(x_{k-2}, x_{k+2})}^{\mu_\varepsilon, u_-, u_+} (\mathbf{1}_{\mathcal{A}_k^\odot})} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_Y). \quad (6.80)$$

It now remains to estimate the ratio of expectations. Recalling (6.78), it suffices to bound the numerator. For this purpose, we remark that (4.25) yields that for any  $\gamma > 0$  and for  $\delta > 0$  sufficiently small, we have

$$\mathbb{E}_{(x_{k-2}, x_{k+2})}^{\mu_\varepsilon, u_-, u_+} (\mathbf{1}_{\mathcal{A}_{w,k}^\odot}) \leq \exp \left( -\frac{1}{\varepsilon} (c_0 - \gamma) \right)$$

(where, as usual, we have redefined  $\gamma$  by a factor of two). Substituting these upper and lower bounds, (6.80) improves to (6.79), and the proof of Lemma 5.4 is complete.  $\square$

*Proof of Lemma 5.5.* We will show (5.8). The proof of (5.9) is similar. We can assume that the interval  $J_{y,-}^\varepsilon$  is contained in  $[-L_\varepsilon, L_\varepsilon]$ ; if it is not, the proof becomes even simpler.

Given the bound (5.4) on  $|\mathcal{I}_-^\varepsilon|$ , it is clearly sufficient to prove that for any fixed  $k \in \mathcal{I}_-^\varepsilon$ , we have

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in \mathcal{J}_{y,\varepsilon} \cap \mathbb{C}\mathcal{A}_{3,y}^- : |u(x_k) + 1| \geq \frac{1}{2} \right) \\ \leq \exp \left( -\frac{3c_1}{4\varepsilon} \right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{y,\varepsilon}). \end{aligned} \quad (6.81)$$

This in turn will follow trivially from

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( u \in \mathcal{J}_{y,\varepsilon} : |u(x_k) + 1| \geq \frac{1}{2}, u \leq 1 - \delta \text{ on } [x_{k-1}, x_{k+1}] \right) \\ \leq \exp \left( -\frac{3c_1}{4\varepsilon} \right) \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}). \end{aligned} \quad (6.82)$$

In order to establish (6.82), we again introduce a decomposition. This time we define the sets

$$\begin{aligned} \mathcal{A}_k^\ominus &:= \{u : |u(x_j)| \leq M \text{ for all } j \leq k-2\}, \\ \mathcal{A}_k^\oplus &:= \{u : |u(x_j)| \leq M \text{ for all } j \geq k+2, \\ &\quad \text{at least one } \delta^- \text{ up layer } \leq 2\ell \text{ in } J_{y,\varepsilon}\}, \\ \mathcal{A}_k^\odot &:= \{u : |u(x_j)| \leq M \text{ for } j = k-2, k-1, \dots, k+2, \\ &\quad u \leq 1 - \delta \text{ on } [x_{k-1}, x_{k+1}]\}, \\ \mathcal{A}_{1/2,k}^\odot &:= \left\{ u \in \mathcal{A}_k^\odot : |u(x_k) + 1| \geq \frac{1}{2} \right\}. \end{aligned}$$

The set on the left-hand side of (6.82) can be written as  $\mathcal{A}_k^\ominus \cap \mathcal{A}_{1/2,k}^\odot \cap \mathcal{A}_k^\oplus$ , and we have the containment

$$\mathcal{A}_k^\ominus \cap \mathcal{A}_k^\odot \cap \mathcal{A}_k^\oplus \subseteq \mathcal{J}_{y,\varepsilon},$$

so that applying the Markov property from Lemma 3.2 leads to

$$\begin{aligned} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{A}_k^\ominus \cap \mathcal{A}_{1/2,k}^\odot \cap \mathcal{A}_k^\oplus) \\ \leq \sup_{u \pm \in [-M, 1-\delta]} \frac{\mathbb{E}^{\mu_\varepsilon, u_-, u_+}_{(x_{k-2}, x_{k+2})} (\mathbf{1}_{\mathcal{A}_{1/2,k}^\odot}(u))}{\mathbb{E}^{\mu_\varepsilon, u_-, u_+}_{(x_{k-2}, x_{k+2})} (\mathbf{1}_{\mathcal{A}_k^\odot}(u))} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} (\mathcal{J}_{y,\varepsilon}). \end{aligned}$$

Therefore, to show the desired estimate (6.82), it is sufficient to establish

$$\sup_{u \pm \in [-M, M]} \frac{\mathbb{E}^{\mu_\varepsilon, u_-, u_+}_{(x_{k-2}, x_{k+2})} (\mathbf{1}_{\mathcal{A}_{1/2,k}^\odot}(u))}{\mathbb{E}^{\mu_\varepsilon, u_-, u_+}_{(x_{k-2}, x_{k+2})} (\mathbf{1}_{\mathcal{A}_k^\odot}(u))} \leq \exp \left( -\frac{3c_1}{4\varepsilon} \right). \quad (6.83)$$

To get a lower bound for the denominator, we will as usual use the large deviation lower bound from Proposition 3.5. For this, we note that

$$\begin{aligned} \mathcal{A}_k^\odot &= B(\mathcal{A}, \delta) \\ \text{where } \mathcal{A} &= \{u : |u(x_j)| \leq M - \delta \text{ for } j = k-2, \dots, k+2, \\ &\quad u \leq 1 - 2\delta \text{ on } [x_{k-1}, x_{k+1}]\}. \end{aligned}$$

Therefore, the large deviation bound gives that for any  $\gamma > 0$  and for  $\varepsilon$  small enough

$$\mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_-, u_+} (\mathcal{A}_k^\odot) \geq \exp \left( -\frac{1}{\varepsilon} (\Delta E(\mathcal{A}) + \gamma) \right). \quad (6.84)$$

To get an upper bound for the numerator of (6.83), on the other hand, we will use the large deviation upper bound from Proposition 3.4. For this, we observe that

the closed  $\delta/2$  ball around  $\mathcal{A}_{1/2,k}^\odot$  is the set

$$\tilde{\mathcal{A}} := \left\{ u : |u(x_k)| \leq M + \delta, \text{ for } j = k-2, \dots, k+2, \right. \\ \left. u \leq 1 - \delta/2 \text{ on } [x_{k-1}, x_{k+1}], |u(x_k) + 1| \geq \frac{1-\delta}{2} \right\},$$

so that the large deviation bound gives

$$\mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_-, u_+} \left( \mathcal{A}_{1/2,k}^\odot \right) \leq \exp \left( -\frac{1}{\varepsilon} (\Delta E(\tilde{\mathcal{A}}) + \gamma) \right). \quad (6.85)$$

We substitute (6.84) and (6.85) into the ratio on the left-hand side of (6.83) and observe that there is a cancellation of the second factor in the energy difference (see equation (3.13)):

$$\frac{\mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_-, u_+} \left( \mathcal{A}_{1/2,k}^\odot \right)}{\mu_{\varepsilon, (x_{k-2}, x_{k+2})}^{u_-, u_+} \left( \mathcal{A}_k^\odot \right)} \leq \exp \left( -\frac{1}{\varepsilon} \left( \inf_{u \in \tilde{\mathcal{A}}} E(u) - \inf_{u \in \mathcal{A}} E(u) - \gamma \right) \right).$$

Hence, the final ingredient that we need is the following energetic fact.

**Lemma 6.7.** *There exists  $C < \infty$  such that for any  $M$  large enough and  $\delta > 0$  small enough, there exists  $\ell_* < \infty$  with the following property. Let  $\ell \geq \ell_*$  and consider the boundary conditions  $u_\pm \in [-M, M]$ . Define the sets  $\mathcal{A}$  and  $\tilde{\mathcal{A}}$  as*

$$\mathcal{A} := \{ u : |u(x)| \leq M - \delta \text{ for } x = -2\ell, -\ell, \dots, 2\ell, \\ u \leq 1 - 2\delta \text{ on } [-\ell, \ell] \}, \\ \tilde{\mathcal{A}} := \left\{ u : |u(x)| \leq M + \delta \text{ for } x = -2\ell, -\ell, \dots, 2\ell, \right. \\ \left. u \leq 1 - \delta/2 \text{ on } [-\ell, \ell], |u(0) + 1| \geq \frac{1-\delta}{2} \right\}.$$

Then there holds

$$\inf_{u \in \tilde{\mathcal{A}}} E_{(x_{k-2}, x_{k+2})}(u) - \inf_{u \in \mathcal{A}} E_{(x_{k-2}, x_{k+2})}(u) \geq c_1 - C\delta,$$

where

$$c_1 := 2 \min \left\{ \int_{-1}^{-1/2} \sqrt{2V(s)} ds, \int_{-3/2}^{-1} \sqrt{2V(s)} ds \right\}. \quad (6.86)$$

This lemma is virtually identical to Lemma 6.6. The principal difference is that here the excursion from  $-1$  is only of magnitude  $1/2$ . This changes only the leading order cost (from  $c_0$  to  $c_1$ ). We omit the proof of the lemma.  $\square$

*Proof of Lemma 5.6.* We will prove only (5.10), the proof of (5.11) being essentially the same. We will always assume that the left endpoint of the interval  $J_{y,-}^\varepsilon$  is greater than or equal to  $-L_\varepsilon$  (since otherwise the boundary condition at  $-L_\varepsilon$  trivially implies the result).

Notice that the set of paths  $u \in \mathcal{J}_{y,\varepsilon}$  that do not hit  $-1$  in  $J_{y,-}^\varepsilon$  is contained in the following two sets

- The set of paths (a) in  $\mathcal{A}_{y,3}^-$  (extra  $\delta^-$  layers: recall (5.5)) or (b) without extra layers but more than  $1/2$  away from  $-1$  at a gridpoint for some  $k$  in  $\mathcal{I}_-^\varepsilon$ .
- The set  $\mathcal{A}_{y,4}^-$  of paths in  $\mathcal{J}_{y,\varepsilon}$  that are within  $1/2$  of  $-1$  at all gridpoints with  $k \in \mathcal{I}_-^\varepsilon$  but do not hit  $-1$  in  $J_{y,-}^\varepsilon$ .

Hence, because of the bounds already established in Lemmas 5.4 and 5.5, we will be done as soon as we show

$$\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{A}_{y,4}^-) \leq \lambda^{\bar{K}_\varepsilon} \mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1}(\mathcal{J}_{y,\varepsilon}). \quad (6.87)$$

We remark for reference below that we may assume  $M \geq 3/2$  so that  $|u(x_k) - 1| \leq 1/2$  implies  $|u(x_k)| \leq M$ .

The interval  $J_{y,-}^\varepsilon$  can naturally be divided up into  $\bar{K}_\varepsilon$  subintervals of length  $\ell(2K_\varepsilon + 1)$ . We set

$$\bar{k}_j := k_{y,-}^\varepsilon + j(2K_\varepsilon + 1) \text{ for } j = 0, \dots, \bar{K}_\varepsilon, \text{ and } \bar{I}_j := [x_{\bar{k}_j}, x_{\bar{k}_{j+1}}] \text{ for } j \leq \bar{K}_\varepsilon - 1.$$

We want to use the Markov property and then apply Lemma 5.2 on these subintervals. Therefore, as usual, we introduce some sets for a decomposition.

$$\begin{aligned} \mathcal{A}^\ominus &:= \{u : |u(x_k)| \leq M \text{ for } k \leq k_{y,-}^\varepsilon\}, \\ \mathcal{A}^\oplus &:= \{u : |u(x_k)| \leq M \text{ for } k \geq k_{y,+}^\varepsilon, \delta^- \text{ up layer} \leq 2\ell \text{ in } J_{y,\varepsilon}\}, \\ \mathcal{A}_{bc}^\odot &:= \{u : |u(x_{\bar{k}_j}) - 1| \leq \frac{1}{2} \text{ for } j = 0, \dots, \bar{K}_\varepsilon\}, \\ \mathcal{A}_j^\odot &:= \{u : |u(x_k) - 1| \leq \frac{1}{2} \text{ for } x_k \in \bar{I}_j\}, \\ \mathcal{A}_{-1,j}^\odot &:= \{u \in \mathcal{A}_j^\odot : \text{no hitting of } -1 \text{ in } \bar{I}_j\}. \end{aligned}$$

We now write  $\mathcal{A}_{y,4}^-$  as the intersection

$$\mathcal{A}_{y,4}^- = \mathcal{A}^\ominus \cap \mathcal{A}^\oplus \cap \mathcal{A}_{bc}^\odot \cap \left( \bigcap_{j=0}^{\bar{K}_\varepsilon-1} \mathcal{A}_{-1,j}^\odot \right), \quad (6.88)$$

and apply the Markov property (Lemma 3.2)  $\bar{K}_\varepsilon$  times to deduce

$$\begin{aligned} &\mu_{\varepsilon,(-L_\varepsilon,L_\varepsilon)}^{-1,1} \left( \mathcal{A}^\ominus \cap \mathcal{A}^\oplus \cap \mathcal{A}_{bc}^\odot \cap \left( \bigcap_{j=0}^{\bar{K}_\varepsilon-1} \mathcal{A}_{-1,j}^\odot \right) \right) \\ &= \mathbb{E}_{(-L_\varepsilon,L_\varepsilon)}^{\mu_\varepsilon, -1, 1} \left( \mathbf{1}_{\mathcal{A}^\ominus}(u) \mathbf{1}_{\mathcal{A}^\oplus}(u) \mathbf{1}_{\mathcal{A}_{bc}^\odot}(u) \prod_{j=0}^{\bar{K}_\varepsilon-1} \mathbb{E}_{(x_{\bar{k}_j}, x_{\bar{k}_{j+1}})}^{\mu_\varepsilon, \mathbf{u}} (\mathbf{1}_{\mathcal{A}_{-1,j}^\odot}(u)) \right). \end{aligned} \quad (6.89)$$

According to Lemma 5.2, we have

$$\mathbb{E}_{(x_{\bar{k}_j}, x_{\bar{k}_{j+1}})}^{\mu_\varepsilon, \mathbf{u}} (\mathbf{1}_{\mathcal{A}_{-1,j}^\odot}(u)) \leq \lambda \mathbb{E}_{(x_{\bar{k}_j}, x_{\bar{k}_{j+1}})}^{\mu_\varepsilon, \mathbf{u}} (\mathbf{1}_{\mathcal{A}_j^\odot}(u)),$$



uniformly over all paths  $\mathbf{u}$  that satisfy  $\mathbf{u}(x_{\bar{k}_j}), \mathbf{u}(x_{\bar{k}_{j+1}}) \in [-3/2, -1/2]$ . We insert this bound into (6.89) and then use the Markov property once more to recover

$$\begin{aligned}
& \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} \left( \mathcal{A}^\ominus \cap \mathcal{A}^\oplus \cap \mathcal{A}_{bc}^\odot \cap \left( \bigcap_{j=0}^{\bar{K}_\varepsilon-1} \mathcal{A}_{-1,j}^\odot \right) \right) \\
& \leq \lambda^{\bar{K}_\varepsilon} \mathbb{E}_{(-L_\varepsilon, L_\varepsilon)}^{\mu_\varepsilon, -1,1} \left( \mathbf{1}_{\mathcal{A}^\ominus}(u) \mathbf{1}_{\mathcal{A}^\oplus}(u) \mathbf{1}_{\mathcal{A}_{bc}^\odot}(u) \prod_{j=0}^{\bar{K}_\varepsilon-1} \mathbb{E}_{(x_{\bar{k}_j}, x_{\bar{k}_{j+1}})}^{\mu_\varepsilon, \mathbf{u}} (\mathbf{1}_{\mathcal{A}_j^\odot}(u)) \right) \\
& = \lambda^{\bar{K}_\varepsilon} \mu_{\varepsilon, (-L_\varepsilon, L_\varepsilon)}^{-1,1} \left( \mathcal{A}^\ominus \cap \mathcal{A}^\oplus \cap \mathcal{A}_{bc}^\odot \cap \left( \bigcap_{j=0}^{\bar{K}_\varepsilon-1} \mathcal{A}_j^\odot \right) \right). \tag{6.90}
\end{aligned}$$

Since

$$\mathcal{A}^\ominus \cap \mathcal{A}^\oplus \cap \mathcal{A}_{bc}^\odot \cap \left( \bigcap_{j=0}^{\bar{K}_\varepsilon-1} \mathcal{A}_j^\odot \right) \subseteq \mathcal{J}_{y,\varepsilon},$$

the combination of (6.88) and (6.90) completes the proof of (6.87).  $\square$

#### ACKNOWLEDGEMENTS

We thank the Max-Planck Institute for Mathematics in the Sciences in Leipzig, where we had the pleasure of working jointly on this project.

The second author would like to thank Volker Betz for explaining to him some background about Schrödinger operators. He would also like to thank Martin Hairer and Andrew Stuart for many discussions about this project and related topics.

The third author would like to thank Eric Vanden-Eijnden for insightful discussions related to ideas developed in this project.

Hendrik Weber was partially supported by ERC grant *AMSTAT – Problems at the Applied Mathematics and Statistics Interface* and by a Philip Leverhulme prize. Maria G. Westdickenberg was partially supported as an Alfred P. Sloan Research Fellow and by the National Science Foundation under Grant No. DMS-0955051.

#### REFERENCES

- [A89] S. ARRHENIUS. Ueber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Z. Phys. Chem.* **4**, (1889), 226–248.
- [BBM10] F. BARRET, A. BOVIER, and S. MÉLÉARD. Uniform estimates for metastable transition times in a coupled bistable system. *Preprint* (2010).
- [B12] F. BARRET. Sharp asymptotics of metastable transition times for one dimensional SPDEs. *Preprint* (2012).
- [BG12] N. BERGLUND and B. GENTZ. Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond. *Preprint*, 2012.
- [BBB08a] L. BERTINI, S. BRASSESCO, and P. BUTTÀ. Soft and hard wall in a stochastic reaction diffusion equation. *Arch. Ration. Mech. Anal.* **190**, no. 2, (2008), 307–345.
- [BBB08b] L. BERTINI, S. BRASSESCO, and P. BUTTÀ. Dobrushin states in the  $\phi_1^4$  model. *Arch. Ration. Mech. Anal.* **190**, no. 3, (2008), 477–516.
- [Bog98] V. I. BOGACHEV. *Gaussian measures*, vol. 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.
- [BEGK04] A. BOVIER, M. ECKHOFF, V. GAYRARD, and M. KLEIN. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**, no. 4, (2004), 399–424.

- [BGK05] A. BOVIER, V. GAYRARD, and M. KLEIN. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc.* **7**, no. 1, (2005), 69–99.
- [BDMP95] S. BRASSESCO, A. DE MASI, and E. PRESUTTI. Brownian fluctuations of the interface in the D=1 Ginzburg-Landau equation with noise. *Ann. Inst. H. Poincaré Probab. Statist.* **31**, no. 1, (1995), 81–118.
- [B93] H. B. BRAUN. Thermally activated magnetization reversal in elongated ferromagnetic particles. *Phys. Rev. Lett.* **71**, no. 21, (1993), 3557–3560.
- [COP93] M. CASSANDRO, E. ORLANDI, and E. PRESUTTI. Interfaces and typical Gibbs configurations for one-dimensional Kac potentials. *Probab. Theory Related Fields* **96**, no. 1, (1993), 57–96.
- [dH00] F. DEN HOLLANDER. *Large deviations*, vol. 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.
- [Ei10] A. EINSTEIN. Theorie der Opaleszenz von homogenen Flüssigkeitsgemischen in der Nähe des kritischen Zustandes. *Ann. Physik* **33**, (1910), 1275–1298.
- [Ey35] H. EYRING. The activated complex in chemical reactions. *J. Chem. Phys.* **3**, (1935), 107–115.
- [Fa27] L. FARKAS. Keimbildungsgeschwindigkeit in übersättigten Dämpfen. *Z. Phys. Chem.* **125**, (1927), 236–242.
- [FJ82] W. G. FARIS and G. JONA-LASINIO. Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A* **15**, no. 10, (1982), 3025–3055.
- [Fe48] R. P. FEYNMAN. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**, no. 2, (1948), 367–387.
- [Fr88] M. I. FREIDLIN. Random perturbations of reaction-diffusion equations: the quasi-deterministic approximation. *Trans. Amer. Math. Soc.* **305**, no. 2, (1988), 665–697.
- [FW98] M. I. FREIDLIN and A. D. WENTZELL. *Random perturbations of dynamical systems*, vol. 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second ed., 1998. Translated from the 1979 Russian original by Joseph Szücs.
- [Fu95] T. FUNAKI. The scaling limit for a stochastic PDE and the separation of phases. *Probab. Theory Related Fields* **102**, no. 2, (1995), 221–288.
- [GLE41] S. GLASSTONE, K. J. LAIDLER, and H. EYRING. *The theory of rate processes*. McGraw-Hill, New York, 1941.
- [Hai09] M. HAIRER. An Introduction to Stochastic PDEs. *ArXiv e-prints*.
- [HTB90] P. HÄNGGI, P. TALKNER, and M. BORKOVEC. Reaction rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, no. 2, (1990), 251–341.
- [Ki74] YU. I. KIFER. Nekotorye rezul'taty, kasayushchiesya malykh sluchaynykh vozmushchenii dinamicheskikh sistem. *Teor. Veroyatnost. i Primenen.* **19**, no. 2, (1974), 514–532. English translation of title: Certain results concerning small random perturbations of dynamical systems.
- [Kr40] H. A. KRAMERS. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, (1940), 284–304.
- [LS61] R. LANDAUER and J. A. SWANSON. Frequency factors in the thermally activated process. *Phys. Rev.* **121**, (1961), 1668–1674.
- [L69] J. S. LANGER. Statistical theory of the decay of metastable states. *Ann. Phys.* **54**, (1969), 258–275.
- [MS77] B. J. MATKOWSKY and Z. SCHUSS. The exit problem for randomly perturbed dynamical systems. *SIAM J. Appl. Math.* **33**, (1977), 365–382.
- [MM77] L. MODICA and S. MORTOLA. Un esempio di  $\Gamma^-$ -convergenza. *Boll. Un. Mat. Ital. B (5)* **14**, no. 1, (1977), 285–299.
- [MoMc89] F. MOSS and P. V. E. MCCLINTOCK. *Noise in Nonlinear Dynamical Systems*. Cambridge University Press, Cambridge, 1989.
- [OM53] L. ONSAGER and S. MACHLUP. Fluctuations and irreversible processes. *Phys. Rev.* **91**, no. 6, (1953), 1505–1512.
- [PAV33] L. S. PONTYAGIN, A. A. ANDRONOV, and A. A. VITT. O statisticheskoy rassmotrenii dinamicheskikh sistem. *Zh. Eksper. Teoret. Fiz.* **3**, no. 3, (1933), 165–180.
- [RY99] D. REVUZ, and M. YOR. Continuous martingales and Brownian motion. Third edition. *Grundlehren der Mathematischen Wissenschaften*, 293. Springer-Verlag, Berlin, 1999.

- [RVE05] M. G. REZNIKOFF and E. VANDEN-EIJNDEN. Invariant measures of stochastic partial differential equations and conditioned diffusions. *C. R. Math. Acad. Sci. Paris* **340**, no. 4, (2005), 305–308.
- [S79] B. SIMON. *Functional integration and quantum physics*. Academic Press, New York, 1979.
- [S95] M. SUGIURA Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Japan* **47**, no. 4, (1995), 755–788.
- [VW08] E. VANDEN-EIJNDEN and M. G. WESTDICKENBERG. Rare events in stochastic partial differential equations on large spatial domains. *J. Stat. Phys.* **131**, (2008), 1023–1038.
- [VH84] J. H. VAN’T HOFF. in *Etudes de Dynamiques Chimiques*, p. 114. F. Muller and Co., Amsterdam, 1884. Translated by T. Ewan as *Studies in Chemical Dynamics* (London, 1896).
- [Var84] S. R. S. VARADHAN. *Large deviations and applications*, vol. 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.
- [Web10] H. WEBER. Sharp interface limit for invariant measures of a stochastic Allen-Cahn equation. *Comm. Pure Appl. Math.* **63**, no. 8, (2010), 1071–1109.
- [WF70] A. D. WENTZELL AND M. I. FREIDLIN. O malykh sluchainykh vozmuschcheniyakh dinamicheskikh sistem. *Uspekhi Mat. Nauk* **25**, no. 1, (1970), 3-55. English translation of title: On small random perturbations of dynamical systems.
- [W30] N. WIENER. Generalized harmonic analysis. *Acta Math.* **55**, no. 1, (1930), 117–258.
- [Zab89] J. ZABCZYK. Symmetric solutions of semilinear stochastic equations. In *Stochastic partial differential equations and applications, II (Trento, 1988)*, vol. 1390 of *Lecture Notes in Math.*, 237–256. Springer, Berlin, 1989.

FELIX OTTO, MPI FOR MATHEMATICS IN THE SCIENCES  
*E-mail address:* felix.otto@mis.mpg.de

HENDRIK WEBER, UNIVERSITY OF WARWICK  
*E-mail address:* hendrik.weber@warwick.ac.uk

MARIA G. WESTDICKENBERG, RWTH AACHEN UNIVERSITY  
*E-mail address:* maria@math1.rwth-aachen.de