Metastability in a continuous mean-field model
at low temperature and strong interaction

K. Bashiri and G. Menz

November 4, 2019

Abstract

We consider a system of $N \in \mathbb{N}$ mean-field interacting stochastic differential equations
that are driven by a single-site potential of double-well form and by Brownian noise. The
strength of the noise is measured by a small parameter $\varepsilon > 0$ (which we interpret as
the temperature), and we suppose that the strength of the interaction is given by $J > 0$.
Choosing the empirical mean $\left( P : \mathbb{R}^N \to \mathbb{R}, P x = 1/N \sum_i x_i \right)$ as the macroscopic order
parameter for the system, we show that the resulting macroscopic Hamiltonian has two
global minima, one at $-m^*_\varepsilon < 0$ and one at $m^*_\varepsilon > 0$. Following this observation, we are
interested in the average transition time of the system to $P^{-1}(m^*_\varepsilon)$, when the initial config-
uration is drawn according to a probability measure (the so-called last-exit distribution),
which is supported around the hyperplane $P^{-1}(-m^*_\varepsilon)$. Under the assumption of strong
interaction, $J > 1$, the main result is a formula for this transition time, which is remi-
niscent of the celebrated Eyring-Kramers formula (see [14]) up to a multiplicative error
term that tends to 1 as $N \to \infty$ and $\varepsilon \to 0$. The proof is based on the potential-theoretic
approach to metastability.

In the last chapter we add some estimates on the metastable transition time in the
high temperature regime, where $\varepsilon = 1$, and for a large class of single-site potentials.

Key words and phrases. Metastability, Local Cramér theorem, Kramers’ law.
2010 Mathematics Subject Classification. 60F10; 60J60; 60K35; 82C22

Introduction

It is by now well-known that many stochastic systems exhibit a phenomenon called metastab-
ility. A typical situation for this is the following. First, for a relatively long time, the system
is trapped in a state, (the metastable state), which is not the (sole) equilibrium state of the
system. Then, after many unsuccessful attempts and due to local fluctuations, the system
finally makes the transition to the (other) equilibrium state (the stable state). In many cases,
this transition is triggered by the appearance of a critical state, and the metastable and the
stable states are modelled through local minima of a free energy functional or Hamiltonian

---

1Institut für Angewandte Mathematik, Rheinische Friedrich-Wilhelms-Universität, Endenicher Allee 60,
53115 Bonn, Germany. Email: bashiri@iam.uni-bonn.de.
2UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555. E-mail:
gmenz@math.ucla.edu.

K.B. is partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)
- Projektnummer 211504053 - SFB 1060 and by Germany’s Excellence Strategy – GZ 2047/1, projekt-id
390685813 – “Hausdorff Center for Mathematics” at Bonn University.
corresponding to the system. For a more detailed introduction to metastability, we refer to [11, Part I].

In this paper, we are interested in the metastable behaviour of a system of \( N \in \mathbb{N} \) stochastic differential equations given by

\[
 dx_i^{N,\varepsilon}(t) = -\psi'(x_i^{N,\varepsilon}(t)) \, dt - \frac{J}{N} \sum_{j=0}^{N-1} \left( x_i^{N,\varepsilon}(t) - x_j^{N,\varepsilon}(t) \right) \, dt + \sqrt{2\varepsilon} \, dB_i(t),
\]

where \( t \in (0, \infty) \), \( 0 \leq i \leq N-1 \), \( \varepsilon > 0 \), \( B^N = (B_i)_{i=0,\ldots,N-1} \) is an \( N \)-dimensional Brownian motion, \( J > 0 \) and the single-site potential \( \psi : \mathbb{R} \to \mathbb{R} \) is given by \( \psi(z) = \frac{1}{4}z^4 - \frac{1}{2}z^2 \). We consider the strength \( \varepsilon \) of the Brownian noise as the temperature of the system.

We proceed as follows. First, in order to analyse the system for large \( N \), we choose the empirical mean, \( P : \mathbb{R}^N \to \mathbb{R}, Px = 1/N \sum_{i=0}^{N-1} x_i \), as the macroscopic order parameter. That is, we consider the image of the system under the map \( P \). Then, as a result of an improvement of the well-known Cramér theorem for this setting, which we call local Cramér theorem (see Section 1.2), we obtain a function \( \bar{H}_\varepsilon : \mathbb{R} \to \mathbb{R} \), which we interpret as the macroscopic Hamiltonian of the system. A simple analysis shows that \( \bar{H}_\varepsilon \) admits exactly two global minima at \( -m_\varepsilon^* < 0 \) and \( m_\varepsilon^* > 0 \), and that \( \bar{H}_\varepsilon \) admits a unique local maximum at 0. This fact indicates that our model exhibits metastable behaviour with the two metastable states being the hyperplanes \( P^{-1}(m_\varepsilon^*) \) and \( P^{-1}(-m_\varepsilon^*) \). The goal of this paper is to compute the average transition time to a region around \( P^{-1}(m_\varepsilon^*) \), when the system is initially close to \( P^{-1}(-m_\varepsilon^*) \).

We tackle this goal in two different regimes, the first one being the low-temperature regime, where the strength \( \varepsilon \) of the Brownian noise tends to zero, and the second one being the high-temperature regime, where we set \( \varepsilon = 1 \). We obtain the following results in this paper.

- In Chapter 2 we show that in the low-temperature regime and under the assumption that \( J > 1 \), the average transition time is asymptotically given by a formula, which is of a similar form as the well-known Eyring-Kramers formula (see [14]) up to a multiplicative error term that tends to 1 as \( N \to \infty \) and \( \varepsilon \to 0 \). Such a result is often known as Kramers’ law in the literature. See [5] for a review on such results.

- In Chapter 3 we consider the high-temperature regime, where we only show that, as \( N \to \infty \), the average transition time is confined to an interval \([\alpha e^{N\Delta}, \beta e^{N\Delta}]\), where \( \Delta = \bar{H}_\varepsilon(0) - \bar{H}_\varepsilon(-m_\varepsilon^*) \), and \( 0 < \alpha < \beta < \infty \) are independent of \( N \). This result still holds true if we replace \( \psi \) by a large class of single-site potentials.

Our proofs are based on the potential-theoretic approach to metastability, which was initiated in the seminal papers [12], [13] and [14]. Here, one uses tools from potential theory to tackle metastability. In particular, one obtains that the average transition time to the (other) equilibrium state can be expressed in terms of quantities from electric networks. More precisely, in terms of capacities, for which powerful variational principles are known. Hence, the computation of sharp estimates basically reduces to an appropriate choice of test functions in those variational principles. The reader is referred to the monograph [11] for an extensive treatment of this approach.

We conclude this introduction with a short remark on the historical background of metastability results in high-dimensional diffusion models. In the papers [2], [3], [6], and [8], Kramers’ law has been shown for systems of \( N \) nearest-neighbour interacting stochastic differential
equations in low temperature. These models are considered as $N$-dimensional approximations of stochastic partial differential equations. A similar setting was studied in [7], where, instead of the potential-theoretic approach, the so-called path-wise approach to metastability was used. This approach, initiated in [15], is motivated by the Freidlin-Wentzell theory, and is based on large deviation estimates. We refer to the monograph [27] for a comprehensive introduction to this approach. In this approach, the asymptotic behaviour of the average transition time is computed up to logarithmic equivalence.

For mean-field interacting systems in the high-temperature regime (i.e. for the setting from Chapter 3 in this paper), the asymptotic behaviour, up to logarithmic equivalence, of the average transition has been stated without proof in [16, Theorem 4]. The rough estimates from Chapter 3 provide a slightly improved version of this conjecture under different initial conditions (see Section 1.4 for more details).

Outline of the paper. In Chapter 1, we introduce the model, formulate the main results, and sketch the main ideas of this paper. In Chapter 2, we provide the full details for the proof of Kramers’ law in the low-temperature regime and under the assumption that $J > 1$, and in Chapter 3 we compute estimates on the average transition time in the high-temperature regime. Finally, in the appendix, we state some general properties of Legendre transforms, compute certain asymptotic integrals by using Laplace’s method, and provide the proof of the local Cramér theorem in a general form, which is the key ingredient for our proofs in this paper.

Notation.

- $\mathcal{P}(Y)$ denotes the space of Borel probability measures on the topological space $Y$.
- For $\mu \in \mathcal{M}_1(Y)$ and a Borel map $f : Y \to \bar{Y}$, $f\#\mu$ is the image measure of $\mu$ by $f$.
- In this paper, $x$ is always an element of $\mathbb{R}^N$ for $N \in \mathbb{N}$, and its components are denoted by $x_i$, $z$ and $m$ are always elements of $\mathbb{R}$.
- Let $K \subset \mathbb{R}^\ell$ for some $\ell \in \mathbb{N}$, and let $f : \mathbb{R}^2 \to [0, \infty)$. In this paper, $O_K(f(\varepsilon,N))$ always stands for a function, whose absolute value is bounded by $f$ uniformly in $K$. That is, $O_K(f(\varepsilon,N)) = R_K(m,\varepsilon,N)$ for some function $R_K : \mathbb{R}^\ell \times \mathbb{R} \times \mathbb{R} \to [0, \infty)$ such that there exists $C_K > 0$ such that for all $m \in K$, for $\varepsilon > 0$ small enough, and for $N \in \mathbb{N}$ large enough, $|R_K(m,\varepsilon,N)| \leq C_K f(\varepsilon,N)$. If we even have that $c_K f(\varepsilon,N) \leq |R_K(m,\varepsilon,N)| \leq C_K f(\varepsilon,N)$ for some $c_K > 0$, we write $\Omega_K(f(\varepsilon,N))$ instead of $O_K(f(\varepsilon,N))$.
- Similarly, $O(f(\varepsilon,N))$ stands for a function $R : \mathbb{R}^2 \to [0, \infty)$ such that $O(f(\varepsilon,N)) = R(\varepsilon,N)$ and there exists a constant $C' > 0$ such that $|R(\varepsilon,N)| \leq C f(\varepsilon,N)$. Finally, we define $\Omega(f(\varepsilon,N))$ analogously as $\Omega_K(f(\varepsilon,N))$.
- Let $(S,d)$ be a metric space, $\rho > 0$ and $s \in S$. Then, define $B_\rho(s) = \{r \in S \mid d(s,r) < \rho\}$.

1 The model and the results

This chapter is organized as follows. In Section 1.1, we define the microscopic model. Then, in Section 1.2, we introduce the macroscopic order parameter, and collect some result on the energy landscape of the model under this order parameter. In Section 1.3 and 1.4, we formulate the two main results of this paper. For the sake of comprehensibility, in this
introduction, we only provide rough formulations of the setting and the main results. For the full details, we refer to the chapters 2 and 3.

1.1 The microscopic model

We consider a system of $N$ stochastic differential equations defined by

$$dx_i^{N,\varepsilon}(t) = -\psi'(x_i^{N,\varepsilon}(t)) \, dt - \frac{J}{N} \sum_{j=0}^{N-1} (x_i^{N,\varepsilon}(t) - x_j^{N,\varepsilon}(t)) \, dt + \sqrt{2\varepsilon} \, dB_i(t), \quad (1.1)$$

where $t \in (0, \infty)$, $0 \leq i \leq N - 1$, $\varepsilon > 0$, $B^N = (B_i)_{i=0,\ldots,N-1}$ is an $N$-dimensional Brownian motion, $J > 0$ and the single-site potential $\psi : \mathbb{R} \to \mathbb{R}$ is given by

$$\psi(z) = \frac{1}{4} z^4 - \frac{1}{2} z^2. \quad (1.2)$$

The Gibbs measure $\mu^{N,\varepsilon} \in \mathcal{P}(\mathbb{R}^N)$ corresponding to this model has the form

$$\mu^{N,\varepsilon}(dx) = \frac{1}{Z_{\mu^{N,\varepsilon}}} e^{-H^{N,\varepsilon}(x)} dx, \quad (1.3)$$

where $Z_{\mu^{N,\varepsilon}}$ is a normalization constant, and, for $x = (x_i)_{i=0,\ldots,N-1} \in \mathbb{R}^N$, the microscopic Hamiltonian $H^{N,\varepsilon} : \mathbb{R}^N \to \mathbb{R}$ is defined by

$$H^{N,\varepsilon}(x) = \frac{1}{\varepsilon} \sum_{i=0}^{N-1} \psi(x_i) + \frac{J}{4N} \sum_{i,j=0}^{N-1} (x_i - x_j)^2. \quad (1.4)$$

For $t \in (0, \infty)$, let $x^{N,\varepsilon}(t) = (x_0^{N,\varepsilon}(t), \ldots, x_{N-1}^{N,\varepsilon}(t))$. It is well-known that $\mu^{N,\varepsilon}$ is the unique stationary measure of the process $(x^{N,\varepsilon}(t))_{t \in (0,\infty)}$.

1.2 The macroscopic variables and the macroscopic energy landscape

The empirical mean $P : \mathbb{R}^N \to \mathbb{R}$ is defined by

$$P \, x = \frac{1}{N} \sum_{i=0}^{N-1} x_i. \quad (1.5)$$

This operator will act as the order parameter for our microscopic system. That is, in order to analyse the process $(x^{N,\varepsilon}(t))_{t \in (0,\infty)}$ for large $N$, we study the image of this process under the map $P$. Therefore, intuitively, $\mu^{N,\varepsilon} := P_#\mu^{N,\varepsilon}$ describes the (long-time) macroscopic behaviour of our microscopic model, and it will be crucial to study the asymptotic behaviour of this measure.

In fact, in Proposition 2.1, we show that, for $\varepsilon$ small enough and for any compact set $K \subset \mathbb{R}$,

$$\tilde{\mu}^{N,\varepsilon}(dm) = e^{-NH_{\varepsilon}(m)} \sqrt{\frac{\varphi''(m)}{2\pi}} \, dm \left( 1 + O_K \left( \frac{1}{\sqrt{N}} \right) \right), \quad (1.6)$$

4
where $\varphi : \mathbb{R} \to \mathbb{R}$ is the so-called Cramér transform of the Gibbs measure with respect to the single-site potential (or more precisely with respect to the modified single-site potential defined in (2.2)) and is defined in (2.9), and $\bar{H}_\varepsilon : \mathbb{R} \to \mathbb{R}$ is defined by

$$\bar{H}_\varepsilon(z) = \varphi_\varepsilon(z) - \frac{1}{\varepsilon} \frac{J}{2} z^2. \quad (1.7)$$

Since $\bar{\mu}_{N,\varepsilon}$ is the law of the empirical mean of a sequence of random variables, (1.6) can be seen as an improvement of the well-known Cramér theorem (cf. [18, 6.1.3]) for this setting. This explains why we call this result local Cramér theorem.

Equation (1.6) shows that, for large $N$ and for $\varepsilon$ small enough, $\bar{\mu}_{N,\varepsilon}$ is very similar to a Gibbs measure with $\bar{H}_\varepsilon$ playing the role of the energy function. Therefore, we consider $\bar{H}_\varepsilon$ as the macroscopic Hamiltonian of the system. This suggests to study the analytic properties of the function $\bar{H}_\varepsilon$. We do this in Lemma 2.2, where we show that, for $\varepsilon$ small enough, $\bar{H}_\varepsilon$ is a symmetric double-well function with two global minima at $-m^*_\varepsilon < 0$ and $m^*_\varepsilon > 0$, and with a local maximum at 0. That is, $\bar{H}_\varepsilon$ is of the form given in Figure 1.

![Figure 1: Form of the graph of the function $\bar{H}_\varepsilon$.](image)

1.3 The Eyring-Kramers formula at low temperature

The fact that the macroscopic free energy $\bar{H}_\varepsilon$ has two global minima at $-m^*_\varepsilon$ and $m^*_\varepsilon$ suggests that our system exhibits metastable behaviour. More precisely, provided that the initial condition of our system is concentrated in a small region around the hyperplane $P^{-1}(-m^*_\varepsilon)$, we expect that the average transition time to hit a small region around $P^{-1}(m^*_\varepsilon)$ fulfils Kramers’ law. This is the content of the main result of this paper, which is formulated in Theorem I. (For a more detailed formulation of the result, we refer to Section 2.2.) In this theorem, we suppose that

$$J > 1. \quad (1.8)$$

The reason for this assumption is that, in this regime, the fluctuations at the macroscopic saddle point 0 can be controlled in an easier way. We explain this in further detail in Remark 2.10. To show the metastable behaviour for the case $J \leq 1$ is the content of future research.

**Theorem I (cf. Theorem 2.7)** Suppose (1.8). Let $T = \inf\{t > 0 | P_{x\varepsilon}^{N,\varepsilon}(t) \geq m^*_\varepsilon - \eta\}$ for some specific $\eta = \Omega(\sqrt{\log(N)/N} \sqrt{\varepsilon \log(\varepsilon^{-1})})$ (see (2.26)). Then, for $\varepsilon$ small enough,
and for $N$ large enough,

$$
E_{\nu_{B^-,B^+}}[T] = \frac{2\pi \sqrt{\varphi''_\varepsilon(-m^*_\varepsilon)} e^{N(\tilde{H}_\varepsilon(0)-\tilde{H}_\varepsilon(-m^*_\varepsilon))}}{\varepsilon \sqrt{\tilde{H}''_\varepsilon(-m^*_\varepsilon)} |\varphi''_\varepsilon(0)| \varphi''_\varepsilon(0)} \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) + O(\varepsilon^2) \right), \quad (1.9)
$$

where $\nu_{B^-,B^+}$ is a probability measure, which is concentrated on the set \( \{ P x = -m^*_\varepsilon + \eta \} \), and is called last-exit biased distribution on $B^-$ (see (2.21) for the definition of $\nu_{B^-,B^+}$ and see (2.27) for the definition of the sets $B^-$ and $B^+$), and $E_{\nu_{B^-,B^+}}[T] := \int E_x[T] \, d\nu_{B^-,B^+}(x)$.

To prove this result in Chapter 2, we proceed as follows.

Before we formulate and prove Theorem I rigorously, we collect in Section 2.1 three important ingredients. More precisely, in Subsection 2.1.1 we state the local Cramér theorem (i.e. (1.6)), which is the key tool in our proof to go from the microscopic variables to the macroscopic ones. Then, in Subsection 2.1.2 we study the analytic properties of the macroscopic Hamiltonian $\tilde{H}_\varepsilon$ and show that its graph is of the form given in Figure 1. And as the third ingredient, we collect in Subsection 2.1.3 the key elements from potential theory that allow us to rewrite the average transition time, $E_{\nu_{B^-,B^+}}[T]$, in terms of quantities from electric networks. Namely, we show that $E_{\nu_{B^-,B^+}}[T]$ is equal to the quotient of the mass of the equilibrium potential and the capacity (see Lemma 2.5 below).

After we collect these ingredients, we formulate the main result of this paper in Section 2.2. The proof of this result follows in three steps. The first step consists of showing the correct upper bound for the capacity in Section 2.3. This is done by using the so-called Dirichlet principle (see Lemma 2.6). Here, we have to choose an appropriate test function and show that the corresponding Dirichlet form is asymptotically equal to the right-hand side of (1.9). In the second step, we compute in Section 2.4 the lower bound on the capacity. To do this, we first make a change of coordinates into Fourier variables. As a consequence of this, we can easily restrict to the macroscopic direction (see Step 2 of the proof of Proposition 2.9). To evaluate the remaining quantity we use large deviation estimates (see Subsection 2.4.2). Finally, we compute in Section 2.5 the asymptotic value of the mass of the equilibrium potential. This follows from applying standard Laplace asympotics, and by exploiting that the graph of $\tilde{H}_\varepsilon$ has the form of a double-well function (cf. Figure 1).

1.4 Rough estimates at high temperature

We also consider in this paper the situation where the microscopic fluctuations in the system do not become negligible. That is, we study the model given by (1.1) with $\varepsilon = 1$. It is not surprising that the methods that we use for the situation in Section 1.3 do not yield the precise Eyring-Kramers formula in the present case. The reason is that in this case, the entropy of the paths matters substantially, i.e. the microscopic fluctuations do not allow to restrict solely to the macroscopic variables under the order parameter $P$. We believe that, in order to obtain the Eyring-Kramers formula, we need to consider the empirical distribution $K : \mathbb{R}^N \to \mathcal{P}(\mathbb{R})$,

$$K x = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{x_i} \quad (1.10)$$

as the order parameter instead of $P$. An heuristic argument for that is the following. Let $(\gamma(t))_{t \in (0,\infty)} \subset \mathcal{P}(\mathcal{P}(\mathbb{R}))$ denote the laws of the random variables $(K(x^N,\varepsilon)(t))_{t \in (0,\infty)}$. Then,
formally, we can write the evolution of \((\gamma(t))_{t \in (0, \infty)}\) as an infinite dimensional stochastic differential equation on the Wasserstein space, i.e.

\[
d\gamma(t) = -\text{Grad}_{\text{Wass}}\mathcal{F}(\gamma(t)) + \frac{1}{\sqrt{N}} \, d\beta(t),
\]

where \(\mathcal{F}\) is the corresponding free energy functional on the Wasserstein space (that is, the laws of the flow \((x^{N,\varepsilon}(t))_{t \in (0, \infty)}\) can be represented as a Wasserstein gradient flow for \(\mathcal{F}\); see [1] or [4]), \(\text{Grad}_{\text{Wass}}\mathcal{F}\) is the gradient of \(\mathcal{F}\) in the Wasserstein space by using the formal Riemannian setting on this space introduced in [28], and \(\beta\) is the corresponding Wasserstein diffusion, which has been introduced in [30]. We refer to [17, (1.8)] for more thoughts on this.

The equation (1.11) heuristically shows that in the mean-field limit as \(N \to \infty\) one is in the same situation as in [14], but in the infinite dimensional Wasserstein space. This suggests that, as in [14], we can expect Kramers’ law to hold for \((\gamma(t))_{t \in (0, \infty)}\).

We note that, in order to show Kramers’ law for \((\gamma(t))_{t \in (0, \infty)}\), there are two approaches to metastability, which seem to be applicable. The first one is based on the characterization of Markov processes as unique solutions of martingale problems (see [24] for an introduction), and the second one is based on the analysis of the corresponding Poisson equation (see [29]).

The rigorous implementation of these thoughts is left for future research.

However, we can still obtain estimates for the mean transition time under the order parameter \(P\). Here we replace \(\psi\) by single-site potentials of the form \(z \mapsto \Psi(z) - \frac{1}{2} \varepsilon z^2\), where \(\Psi : \mathbb{R} \to \mathbb{R}\) is a symmetric and bounded perturbation of a strictly convex function (cf. Assumption 3.1). Moreover, we have to assume that \(J \geq \int_\mathbb{R} e^{-\Psi(z)} \, dz / (\int \varepsilon^2 e^{-\Psi(z)} \, dz)\). This condition is necessary for \(\hat{H}_1\) to be of the form of a double-well function. (Note that the objects \(\varphi_1, \hat{H}_1, \nu_{B_{1-},B_{1+}}, \mathcal{T}\) are defined as in Theorem I but with \(\psi\) replaced by \(z \mapsto \Psi(z) - \frac{1}{2} \varepsilon z^2\) and with \(\varepsilon = 1\).) That is, in the case \(J \leq \int_\mathbb{R} e^{-\Psi(z)} \, dz / (\int \varepsilon^2 e^{-\Psi(z)} \, dz)\), we do not have a metastable behaviour for the system under the order parameter \(P\). This is different than in Theorem I, where we can show that \(\hat{H}_1\) is a double-well function also in the case \(J \leq 1\) (see Lemma 2.2). The main result is the following statement.

**Theorem II (cf. Theorem 3.5)** Suppose Assumption 3.1. Let \(\pm m_1^*\) be the two global minimisers of the macroscopic Hamiltonian \(\hat{H}_1\). Then, for some \(a > 0\), which is independent of \(N\),

\[
\mathbb{E}_{\nu_{B_{1-},B_{1+}}}[\mathcal{T}] \geq \frac{2 \pi \sqrt{\varphi_1''(-m_1^*)} \, e^{N(\hat{H}_1(0) - \hat{H}_1(-m_1^*))}}{\sqrt{\hat{H}_1''(-m_1^*)}|\hat{H}_1''(0)| \phi_1''(0)} \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) \right), \quad \text{and}
\]

\[
\mathbb{E}_{\nu_{B_{1-},B_{1+}}}[\mathcal{T}] \leq (1 + a) \frac{2 \pi \sqrt{\varphi_1''(-m_1^*)} \, e^{N(\hat{H}_1(0) - \hat{H}_1(-m_1^*))}}{\sqrt{\hat{H}_1''(-m_1^*)}|\hat{H}_1''(0)| \phi_1''(0)} \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) \right). \quad (1.12)
\]

The proof of this result is organized in the same way as the proof of Theorem I, and is given in Chapter 3.

Finally, we point out that Theorem II provides a slight improvement of the conjecture given in [16, Theorem 4]. Indeed, the authors of [16] expect that for all \(\delta > 0\), there exists \(N_\delta \in \mathbb{N}\) such that for \(N \geq N_\delta\) the expected transition time is confined to the interval \([e^{N(\Delta - \delta)}, e^{N(\Delta + \delta)}]\), where \(\Delta = \hat{H}_1(0) - \hat{H}_1(-m_1^*)\). Here we have used the simple fact that \(\hat{H}_1(0) - \hat{H}_1(-m_1^*)\) can be written in terms of the free energy functional \(\mathcal{F}\) from (1.11). (We refer to [26, Section IV.2] for more details on this relation.) However, we note that the initial condition in our setting is different than in the conjecture formulated in [16, Theorem 4].
2 The Eyring-Kramers formula at low temperature

In order to simplify the notation, we omit in this chapter the superscripts $N$ and $\varepsilon$. For example, we abbreviate $x = x^{N,\varepsilon}$, $H = H^{N,\varepsilon}$ and $\mu = \mu^{N,\varepsilon}$. Moreover, we rewrite the microscopic Hamiltonian $H$ (see (1.4)) as

$$H(x) = \frac{1}{\varepsilon} \sum_{i=0}^{N-1} \psi_J(x_i) - \frac{1}{\varepsilon} \frac{J}{2N} \sum_{i,j=0}^{N-1} x_i x_j,$$

(2.1)

where the (modified) single-site potential $\psi_J : \mathbb{R} \to \mathbb{R}$ is defined by

$$\psi_J(z) = \psi(z) + \frac{J}{2} z^2 = \frac{1}{4} z^4 + \frac{J-1}{2} z^2.$$

(2.2)

Recall that, for the strength of the interaction part in this model, we assume that $J > 1$.

(2.3)

The outline of this chapter is given after Theorem I in Section 1.3.

2.1 Preliminaries

2.1.1 Local Cramér theorem

This subsection extends the results from [21, Proposition 31] or [25, Section 3]. The goal is to find an asymptotic representation for the measure $\bar{\mu} = P_{\#} \mu$.

The first observation is that we can disintegrate $\mu$ with respect to $\bar{\mu}$ explicitly using the coarea formula ([20, Section 3.4.2]). Indeed, as in [21, p. 306], we obtain that

$$\int_{\mathbb{R}^N} f(x) \, d\mu(x) = \int_{\mathbb{R}} \int_{P^{-1}(m)} f(x) \, d\mu_m(x) \, d\bar{\mu}(m)$$

(2.4)

for all bounded and measurable $f : \mathbb{R}^N \to \mathbb{R}$, where the conditional measures (or fluctuation measures) $\mu_m$ are given by

$$d\mu_m(x) = \mathbb{1}_{P^{-1}(m)}(x) \, e^{-\frac{1}{\varepsilon} \sum_{i=0}^{N-1} \psi_J(x_i)} \, dH^{N-1}(x) \, e^{N\varphi_{N,\varepsilon}(m)},$$

(2.5)

and $\varphi_{N,\varepsilon} : \mathbb{R} \to \mathbb{R}$ is defined by

$$\varphi_{N,\varepsilon}(m) = -\frac{1}{N} \log \int_{P^{-1}(m)} e^{-\frac{1}{\varepsilon} \sum_{i=0}^{N-1} \psi_J(x_i)} \, dH^{N-1}(x).$$

(2.6)

Moreover, for $\bar{\mu}$, we obtain the representation

$$d\bar{\mu}(m) = \frac{1}{Z_{\bar{\mu}}} \, e^{-N\varphi_{N,\varepsilon}(m)+\frac{1}{\varepsilon} N \frac{J}{2} m^2} \, dm$$

(2.7)

for some normalization constant $Z_{\bar{\mu}}$.

It turns out that the asymptotic behaviour of $\bar{\mu}$ will be determined by the Cramér transform $\varphi_{\varepsilon}$ of the measure $e^{-\frac{1}{\varepsilon} \psi_J(z)} \, dz$, which is defined as the Legendre transform of the function

$$\mathbb{R} \ni \sigma \mapsto \varphi^{*}_{\varepsilon}(\sigma) = \log \int_{\mathbb{R}} e^{\sigma z - \frac{1}{\varepsilon} \psi_J(z)} \, dz \in \mathbb{R}.$$

(2.8)
That is,
\[ \varphi_{\varepsilon}(m) = \sup_{\sigma \in \mathbb{R}} (\sigma m - \varphi_{\varepsilon}^*(\sigma)). \] (2.9)

Moreover, for \( \sigma \in \mathbb{R} \), we define the probability measure \( \mu^{\varepsilon,\sigma} \in \mathcal{P}(\mathbb{R}) \) by
\[ d\mu^{\varepsilon,\sigma}(z) = e^{-\int z \psi_{\varepsilon}(z) + \varepsilon \frac{1}{2} \bar{H}_{\varepsilon}(z) \frac{1}{\sqrt{2\pi}} \left( 1 + O_K \left( \frac{1}{\sqrt{N}} \right) \right)} dm, \] (2.10)

\( \mu^{\varepsilon,\sigma} \) is closely related to \( \varphi_{\varepsilon}^* \) and \( \varphi_{\varepsilon} \). This can be seen in Section A in the appendix, where we list several properties of the Cramér transform that are used in this paper.

In the following proposition we state the local Cramér theorem. (Recall that in Section 1.2 we explain why this result is called like that.) Very similar versions of this result are already known in the literature; see for instance [21, Proposition 31] or [25, Section 3]. The main novelty here is that the result is uniform in \( \varepsilon \ll 1 \).

**Proposition 2.1 (Local Cramér theorem)** Suppose (2.3). Let \( K \subset \mathbb{R} \) be compact. Then, there exists \( \varepsilon_K > 0 \) such that, for all \( \varepsilon < \varepsilon_K \) and for all \( m \in K \),
\[ e^{-N\varphi_{\varepsilon,K}(m)} = e^{-N\varphi_{\varepsilon}(m)} \frac{\sqrt{\varphi''_{\varepsilon}(m)}}{\sqrt{2\pi}} \left( 1 + O_K \left( \frac{1}{\sqrt{N}} \right) \right). \] (2.11)

In particular, this implies that for all \( \varepsilon < \varepsilon_K \) and for all \( m \in K \),
\[ d\bar{\mu}(m) = \frac{1}{Z_{\bar{\mu}}} e^{-N\bar{H}_{\varepsilon}(m)} \frac{\sqrt{\varphi''_{\varepsilon}(m)}}{\sqrt{2\pi}} \left( 1 + O_K \left( \frac{1}{\sqrt{N}} \right) \right) dm, \] (2.12)

where
\[ \bar{H}_{\varepsilon}(z) = \varphi_{\varepsilon}(z) - \frac{1}{\varepsilon} \frac{J}{2} z^2. \] (2.13)

**Proof.** We prove this claim in Proposition C.2 in the appendix for a more general setting. \( \square \)

### 2.1.2 Analysis of the energy landscape

Proposition 2.1 indicates that the graph of \( \bar{H}_{\varepsilon} \) determines the macroscopic energy landscape of our system under the order parameter \( P \) (see also Section 1.2 for more comments). This suggests to study the analytic properties of \( \bar{H}_{\varepsilon} \), which is the content of the following lemma.

**Lemma 2.2** Suppose that \( J > 0 \). Then,

(i) \( \lim_{|t| \to \infty} \frac{1}{|t|} \varphi_{\varepsilon}(t) = \infty \), \( \lim_{|t| \to \infty} \frac{1}{|t|} \bar{H}_{\varepsilon}(t) = \infty \), and

(ii) for all \( \varepsilon > 0 \) small enough, \( \bar{H}_{\varepsilon} \) has exactly three critical points located at \(-m_{\varepsilon}^*, 0 \) and \( m_{\varepsilon}^* \) for some \( m_{\varepsilon}^* = 1 + \Omega(\varepsilon) \). Moreover, \( \bar{H}_{\varepsilon}''(0) < 0, \bar{H}_{\varepsilon}''(m_{\varepsilon}^*) > 0 \) and \( \bar{H}_{\varepsilon}''(-m_{\varepsilon}^*) > 0 \). That is, \( \bar{H}_{\varepsilon} \) has a local maximum at 0, and the two global minima of \( \bar{H}_{\varepsilon} \) are located at \( \pm m_{\varepsilon}^* \).
Proof. Part (i) follows from a simple argument, which is based on the fact that $\psi_J$ is super-quadratic at infinity and on H"older’s inequality. For instance, a proof can be found in [26, III.2.6] for a slightly more general setting.

To show part (ii), first note that by Lemma A.1, the condition $\bar{H}_\varepsilon'(m) = 0$ is equivalent to

$$m = (\varphi_\varepsilon^*)' \left( \frac{1}{\varepsilon} J m \right) = \int_{\mathbb{R}} z e^{-\varphi_\varepsilon^* \left( \frac{1}{\varepsilon} J m \right) + \frac{1}{\varepsilon} J zm - \frac{1}{\varepsilon} \psi_J(z)} \, dz.$$  

(2.14)

We know from [22, 3.1 and 3.2] that, for $\varepsilon$ small enough, there exist exactly three solutions $\pm m^*_\varepsilon$ and 0 for (2.14), where $m^*_\varepsilon = 1 + \Omega(\varepsilon)$.

We now show that $\bar{H}_\varepsilon''(0) < 0$ in the case $J > 1$. Using that $\varphi_\varepsilon'(m^*_\varepsilon) = 0$, Lemma A.1 and Corollary B.2 yield that

$$\bar{H}_\varepsilon''(0) = \left( \int_{\mathbb{R}} z^2 \, d\mu_\varepsilon(\varepsilon, \bar{z}) \right)^{-1} - \frac{1}{\varepsilon} J = \frac{J - 1}{\varepsilon} \Omega(\varepsilon) - \frac{1}{\varepsilon} J < 0$$  

(2.15)

for $\varepsilon$ small enough. In the case $J < 1$, we have by standard Laplace asymptotics that for $\varepsilon$ small enough, $\bar{H}_\varepsilon''(0) = \Omega(1) - \frac{1}{\varepsilon} J < 0$. The same result holds also for the case $J = 1$, since $\bar{H}_\varepsilon''(0)$ depends continuously on $J$ (cf. Step 5.3 in the proof of [22, 3.2]).

By the symmetry of $\bar{H}_\varepsilon$, it only remains to show that $\bar{H}_\varepsilon''(m^*_\varepsilon) > 0$. First note that, since $m^*_\varepsilon = 1 + \Omega(\varepsilon)$, for all $J > 0$, the function $z \mapsto \psi_J(z) - Jm^*_\varepsilon z$ admits a unique global minimum at some point $z_\varepsilon = 1 + \Omega(\varepsilon)$. Indeed, in the case $J > 1$, this follows by simply observing that $\psi_J'$ is invertible, and in the case $J \leq 1$, we have to apply Cardano’s formula (see [9, Chapter 1 and 2]). (We omit the details in the latter case, since we do not use the claim of this lemma for the case $J \leq 1$ in the remaining part of this paper.) Then, as above, using Lemma A.1, Corollary B.2 and that $\varphi_\varepsilon'(m^*_\varepsilon) = Jm^*_\varepsilon$ implies that for $\varepsilon$ small enough,

$$\bar{H}_\varepsilon''(m^*_\varepsilon) = \left( \int_{\mathbb{R}} \left( z - \int_{\mathbb{R}} \bar{z} \, d\mu_\varepsilon(\varepsilon, \bar{z}) \right)^2 \, d\mu_\varepsilon(\varepsilon, \bar{z}) \right)^{-1} - \frac{1}{\varepsilon} J$$

$$= \frac{1}{\varepsilon} \psi_J'(z_\varepsilon) \left( 1 + O(\varepsilon \log(\varepsilon^{-1})^2) \right) - \frac{1}{\varepsilon} J$$

$$= \frac{1}{\varepsilon} (3z_\varepsilon^2 - 1) \left( 1 + O(\varepsilon \log(\varepsilon^{-1})^2) \right) > 0,$$

which concludes the proof. \(\square\)

Remark 2.3 In the remaining part of this paper, we suppose that $\varepsilon$ is small enough such that $[-m^*_\varepsilon, m^*_\varepsilon] \subset [-2, 2]$.

2.1.3 Potential-theoretic approach to metastability

In this subsection, we quickly review the key ingredients from the potential-theoretic approach to metastability that we need in our setting. We follow [14, Chapter 2], where all the omitted details can be found.

The generator of the stochastic process $(x(t))_{t \in (0, \infty)}$ introduced in Section 1.1 is given by

$$\mathcal{L} = \varepsilon e^{H} \left( \nabla e^{-H} \nabla \right).$$  

(2.17)

where $H$ is the microscopic Hamiltonian (recall (2.1)). We need the following definitions.
Definition 2.4 Let $A, D \subset \mathbb{R}^N$ be open and regular and such that $A \cap D = \emptyset$ and $(A \cup D)^c$ is connected. For any $B \subset \mathbb{R}^N$, we write $T_B = \inf\{t > 0 \mid x(t) \in B\}$.

(i) The equilibrium potential between $A$ and $D$, $f^*_{A,D}$, is defined as the unique solution to the Dirichlet problem

$$(-\mathcal{L} f)(x) = 0, \quad \text{for } x \in (A \cup D)^c,$$

$$f(x) = 1, \quad \text{for } x \in A,$$

$$f(x) = 0, \quad \text{for } x \in D.$$ 

(2.18)

For $x \in (A \cup D)^c$, we have the probabilistic interpretation that $f^*_{A,D} = \mathbb{P}_x[T_A < T_D]$.

(ii) The equilibrium measure, $e_{A,D}$, is defined as the unique measure on $\partial A$ such that

$$f^*_{A,D}(x) = \int_{\partial A} G_{D^c}(x,y) e_{A,D}(dy) \quad \text{for } x \in (A \cup D)^c,$$ 

(2.19)

where $G_{D^c}$ is the Green function corresponding to $\mathcal{L}$ on $D^c$ (cf. [14, (2.2)]).

(iii) The capacity, $\text{Cap}(A,D)$, of the capacitor $(A,D)$ is defined by

$$\text{Cap}(A,D) = \int_{\partial A} e^{-H(y)} e_{A,D}(dy).$$ 

(2.20)

(iv) The last-exit biased distribution on $A$, $\nu_{A,D}$, is the probability measure on $\partial A$ defined by

$$\nu_{A,D}(dy) = \frac{e^{-H(y)} e_{A,D}(dy)}{\text{Cap}(A,D)}.$$ 

(2.21)

Using these notions, one can rewrite the average hitting time of $B$ in the case that the initial condition is randomly chosen according to the last-exit distribution. This is the content of the following lemma.

Lemma 2.5 Consider the same setting as in Definition 2.4. Then,

$$\mathbb{E}_{\nu_{A,D}}[T_D] := \int_{\partial A} \mathbb{E}_y[T_D] \nu_{A,D}(dy) = \frac{\int_{D^c} f^*_{A,D}(y) e^{-H(y)} dy}{\text{Cap}(A,D)}.$$ 

(2.22)

Proof. The proof can be found in [11, 7.30]. See also [14, (2.27)]. □

As we already mentioned, the main advantage to use Lemma 2.5 is the availability of variational principles for the capacity. In this paper, we use the so-called Dirichlet principle, which is stated in the following lemma.

Lemma 2.6 (Dirichlet principle) Consider the same setting as in Definition 2.4. Let

$$\mathcal{H}_{A,D} = \left\{ f \in H^1(\mathbb{R}^N; e^{-H(x)} dx) \mid f|_A = 1, \ f|_D = 0, \forall x \in \mathbb{R}^N : f(x) \in [0,1], \right\},$$

(2.23)

and define the Dirichlet form on $(A \cup D)^c$, $\mathcal{E}_{(A\cup D)^c}: \mathcal{H}_{A,D} \to [0,\infty]$, by

$$\mathcal{E}_{(A\cup D)^c}(f) = \varepsilon \int_{(A\cup D)^c} |\nabla f(x)|^2 e^{-H(x)} dx \quad \text{for } f \in \mathcal{H}_{A,D}.$$ 

(2.24)

Then,

$$\text{Cap}(A,D) = \inf_{f \in \mathcal{H}_{A,D}} \mathcal{E}_{(A\cup D)^c}(f) = \mathcal{E}_{(A\cup D)^c}(f^*_{A,D}).$$ 

(2.25)

Proof. The proof can be found in [11, 7.33]. See also [14, (2.15)]. □
2.2 The Eyring-Kramers formula

We have now collected all the notions that we need to formulate the main result in this paper. Recall that, under (2.3) and for \( \varepsilon \) small enough, the macroscopic Hamiltonian admits exactly two global minima \( \pm m^* \). We therefore consider the hyperplanes \( P^{-1}(-m^*) \) and \( P^{-1}(m^*) \) as the metastable sets in our system.

The goal in this paper is to use the potential-theoretic setting from Subsection 2.1.3 to compute the average transition time from \( P^{-1}(-m^*) \) to \( P^{-1}(m^*) \) for the stochastic process \( (x(t))_{t \in (0,\infty)} \) introduced in Section 1.1. However, due to technical reasons, we have to modify this goal in two ways.

First, instead of considering \( P^{-1}(-m^*) \) and \( P^{-1}(m^*) \) as the metastable sets, we rather consider \( P^{-1}(-m^* + \eta) \) and \( P^{-1}(m^* - \eta) \), where

\[
\eta = \frac{\sqrt{2}}{\sqrt{NH^{\varepsilon}_e(-m^*)}} \sqrt{\log(N\varepsilon^{-1})}.
\]  

(2.26)

(By using the arguments from Step 3 of the proof of Proposition 2.8, we have that \( \eta = \Omega(\sqrt{\log(N)/N} \sqrt{\varepsilon \log(\varepsilon^{-1})}) \). Heuristically, the reason for this shift is the following. In the proof of our main result, we have to compute the integral in the numerator on the right-hand side of (2.22). Using the disintegration (2.4), Proposition 2.1 and the fact that \( \bar{H}_\varepsilon \) has its global minima at \( -m^* \) and \( m^* \), we see that this integral is concentrated on the sets \( \{ x | Px \in [\pm m^* - \eta, \pm m^* + \eta] \} \). Hence, in order to apply Laplace’s method, we need that the equilibrium potential is equal to 1 or equal to 0 on these sets, respectively.

Second, instead of running the system from some specific point in \( P^{-1}(-m^* + \eta) \), we rather have to initialise our system randomly according to the last-exit biased distribution \( \nu_{B^-} \), where \( B^- \subset \mathbb{R}^N \) are defined by

\[
B^- = \{ x \in \mathbb{R}^N \mid Px \leq -m^* + \eta \} \quad \text{and} \quad B^+ = \{ x \in \mathbb{R}^N \mid Px \geq m^* - \eta \}. \quad \text{(2.27)}
\]

Note that \( \nu_{B^-} \) is a probability measure supported on \( \partial B^- = P^{-1}(-m^* + \eta) \). The main reason for the choice of this initial distribution is that we can exploit the formula (2.22). However, in a finite-dimensional setting, such as in [14], we could also obtain an asymptotic expression for \( \mathbb{E}_y[T_{B^+}] \) for \( y \in \partial B^- \). This is done by using Harnack inequalities. But since these inequalities depend on the dimension of the base space, we are not able to transfer the strategy used in [14] to our high-dimensional setting.

In the following theorem we formulate the first main result of this paper.

**Theorem 2.7** Suppose (2.3), and recall the definition of \( B^- \) and \( B^+ \) in (2.27). Then,

\[
\mathbb{E}_{\nu_{B^-}}[T_{B^+}] = \frac{2\pi}{\varepsilon} \frac{\sqrt{\varphi^{\varepsilon}_e(-m^*)} e^{N(\bar{H}(0) - \bar{H}_e(-m^*))}}{\sqrt{H^{\varepsilon}_e(-m^*)}|\varphi^{\varepsilon}_e(0)|} \left( 1 + O \left( \frac{\log(N)^3}{N} \right) + O \left( \varepsilon^3 \right) \right). \quad \text{(2.28)}
\]

**Proof.** Combining (2.22) with the Propositions 2.8, 2.9 and 2.14 concludes the proof. \( \square \)

2.3 Upper bound on the capacity

**Proposition 2.8** Consider the same setting as in Theorem 2.7. Then, for \( \varepsilon \) small enough,

\[
\text{Cap}(B^-, B^+) \leq \varepsilon \frac{1}{2\pi} e^{-N\bar{H}(0)} \sqrt{|\bar{H}^{\varepsilon}_e(0)|} \sqrt{\varphi^{\varepsilon}_e(0)} \left( 1 + O \left( \frac{\log(N)^3}{N} \right) \right). \quad \text{(2.29)}
\]
Proof. We will obtain the upper bound by using the Dirichlet principle (Lemma 2.6). That is, we introduce a suitable test function and show that the corresponding Dirichlet form is asymptotically given by the right-hand side of (2.29).

**Step 1.** [Choice of the test function $f$.]

Let

$$\rho = \frac{1}{\sqrt{N |H_\mu'(0)|}} \sqrt{\log(N)}$$

and

$$h^*(m) = \frac{\int_{\rho}^\rho \varphi''_e(m) e^{\frac{1}{2} e^{NH_e(m)}} \, dm}{\int_{-\rho}^\rho \varphi''_e(m) e^{\frac{1}{2} e^{NH_e(m)}} \, dm},$$

(2.30)

which is well-defined, since $\varphi_e$ is strictly convex. Then, $h^*$ is the equilibrium potential corresponding to the invariant measure $1_{(-\rho,\rho)}(z) \varphi''_e(m) e^{-NH_e(m)} \, dz$; see [11, Section 7.2.5]. The test function that we use in this proof is given by

$$f(x) = \begin{cases} 
1 & \text{if } Px \leq -\rho, \\
0 & \text{if } Px \geq \rho, \\
h^*(Px) & \text{if } Px \in (\rho, \rho).
\end{cases} \quad (2.31)$$

**Step 2.** [Estimation of the Dirichlet form of $f$.]

Using Lemma 2.6 and (2.4), we have the following upper bound for the capacity.

$$\frac{1}{Z_\mu} \text{Cap}(B^-, B^+) \leq \epsilon \frac{1}{\sqrt{2\pi N}} \sum_{i=0}^{N-1} \left| \partial_i h^* \left( \frac{1}{N} \sum_{i=0}^{N-1} x_i \right) \right|^2 d\mu$$

$$= \epsilon \frac{1}{N} \int_{\{x \in \mathbb{R}^N \mid Px \in (-\rho, \rho)\}} |(h^*)(Px)|^2 \, d\mu$$

$$= \epsilon \frac{1}{N} \int_{-\rho}^\rho \int_{P^{-1}(m)} |(h^')(m)|^2 \, d\mu \, dm \, d\bar{\mu}(m).$$

(2.32)

Applying Proposition 2.1 for $K = [-2, 2]$ and the definition of $h^*$ yields that

$$\frac{1}{Z_\mu} \text{Cap}(B^-, B^+) \leq \frac{1}{\sqrt{2\pi N Z_\mu}} \int_{-\rho}^\rho |(h^')(m)|^2 \sqrt{\varphi''_e(m)} e^{NH_e(m)} \, dm \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right)$$

$$= \frac{\epsilon}{\sqrt{2\pi N Z_\mu}} \left( \int_{-\rho}^\rho \frac{1}{\sqrt{\varphi''_e(m)}} e^{NH_e(m)} \, dm \right)^{-1} \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right).$$

(2.33)

In Step 4 and 5 of this proof we show that for $m \in [-\rho, \rho],$

$$\sqrt{\varphi''_e(m)} = \sqrt{\varphi''_e(0)} \left( 1 + O_K \left( \frac{\sqrt{\log(N)}}{N} \right) \right), \quad \text{and}$$

$$H_e(m) = H_e(0) + \frac{1}{2} m^2 H''_e(0) + O_K \left( \sqrt{\epsilon} \sqrt{\log(N)} \frac{\log(N)}{N} \right).$$

(2.34)

And since, by the coarea formula, $Z_\mu \sqrt{N} = Z_\mu,$ (2.34) and (2.35) imply that

$$\text{Cap}(B^-, B^+) \leq \frac{\epsilon \sqrt{\varphi''_e(0)} e^{-NH_e(0)}}{\sqrt{2\pi N Z_\mu}} \left( \int_{-\rho}^\rho e^{\frac{1}{2} NH''_e(0)} \, dm \right)^{-1} \left( 1 + O \left( \sqrt{\frac{\log(N)}{N}} \right) \right).$$

(2.36)
Combining this with the fact that
\[
\int_{-\rho}^{\rho} e^{\frac{1}{2} N m^2 \rho''(0)} \, dm \geq \sqrt{\frac{2\pi}{N |H''_\rho(0)|}} \left( 1 - e^{\frac{1}{2} N \rho^2 \rho''(0)} \right)^{\frac{1}{2}} = \sqrt{\frac{2\pi}{N |H''_\rho(0)|}} \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right),
\]
(2.37)
concludes the proof of (2.29).

**Step 3.** [Some a priori estimates.]
Before we show (2.34) and (2.35), we collect some a priori estimates. First, we use (A.5) and Lemma C.1 (iii) to see that there exists \( c > 0 \) such that for all \( m \in K \) and \( \varepsilon \) small enough,
\[
|\varphi''_\varepsilon(m)| = \varphi''_\varepsilon(m) = \frac{1}{(\varphi''_\varepsilon)''(\varphi'(m))} \in \left[ \frac{c^{-1}}{\varepsilon}, \frac{c}{\varepsilon} \right].
\]
(2.38)
Moreover, recall that in the proof of Lemma 2.2, we have seen that
\[
|\bar{H}''_\varepsilon(0)| = \frac{1}{\varepsilon(1 + O(\varepsilon))}.
\]
Therefore, for \( \varepsilon \) small enough,
\[
|\bar{H}''_\varepsilon(0)| \geq \frac{1}{4\varepsilon}.
\]
Next, we recall the definition of \( \mu_{\varepsilon, \sigma} \) in (2.10) and use Lemma A.1, (2.38) and Corollary B.2 to see that there exists \( c' > 0 \) such that
\[
|\varphi'''_\varepsilon(m)| = \left| \left( \frac{(\varphi''_\varepsilon)''(\varphi'(m))}{(\varphi''_\varepsilon)''(\varphi'(m))} \right)^3 \right| = \left| \int_{\mathbb{R}} (z - m)^3 \, d\mu_{\varepsilon, \sigma}(\theta m)(z) \right| \varphi''_\varepsilon(m)^3 \in \left[ \frac{(c')^{-1}}{\varepsilon}, \frac{c'}{\varepsilon} \right].
\]
(2.39)

**Step 4.** [Proof of (2.34).]
By Taylor’s formula, we have for some \( \theta \in [0, 1] \),
\[
\sqrt{\varphi''_\varepsilon(m)} = \sqrt{\varphi''_\varepsilon(0)} \left( 1 + m \frac{\varphi''_\varepsilon(\theta m)}{2\sqrt{\varphi''_\varepsilon(0)} \sqrt{\varphi''_\varepsilon(\theta m)}} \right).
\]
(2.40)
Then, by the estimates from Step 3,
\[
\left| m \frac{\varphi''_\varepsilon(\theta m)}{2\sqrt{\varphi''_\varepsilon(0)} \sqrt{\varphi''_\varepsilon(\theta m)}} \right| \leq \frac{1}{2} \rho c c' = \frac{1}{2} \sqrt{\frac{\log(N)}{N}} \frac{c c'}{\sqrt{|H''_\rho(0)|}} \leq c' \sqrt{\frac{\varepsilon \log(N)}{N}}.
\]
(2.41)
In combination with (2.40), this yields (2.34).

**Step 5.** [Proof of (2.35).]
Again by Taylor’s formula, for some \( \theta' \in [0, 1] \),
\[
\bar{H}_\varepsilon(m) = \bar{H}_\varepsilon(0) + \frac{1}{2} m^2 \bar{H}''_\varepsilon(0) + \frac{1}{6} m^3 \bar{H}'''_\varepsilon(\theta' m).
\]
(2.42)
Similarly as in Step 4, we have that
\[
|m^3 \bar{H}'''_\varepsilon(\theta' m)| \leq \rho^3 |\varphi'''_\varepsilon(\theta' m)| \leq \sqrt{\frac{\log(N)}{N}} \frac{1}{\sqrt{|H''_\rho(0)|}} \frac{c'}{\varepsilon} \leq 8 c' \sqrt{\varepsilon} \sqrt{\frac{\log(N)}{N}},
\]
(2.43)
which concludes the proof of (2.35). \( \square \)
2.4 Lower bound on the capacity

**Proposition 2.9** Consider the same setting as in Theorem 2.7. Then, for \( \varepsilon \) small enough,

\[
\Cap(B^-, B^+) \geq \frac{\varepsilon}{2\pi} e^{-NR_0(0)} \sqrt{|H_\varepsilon'(0)|} \sqrt{r_\varepsilon''(0)} \left( 1 + O \left( \sqrt{\frac{\log(N)^3}{N}} \right) + O(\varepsilon^2) \right). \tag{2.44}
\]

Similarly as in the papers [3] and [6], for the proof of Proposition 2.9, we change into the Fourier basis. The advantage is that, in the new variables, the macroscopic order parameter corresponds to the 0-th coordinate (or 0-th mode). Therefore, we can restrict the gradient in the Dirichlet form easily to the macroscopic direction by neglecting all partial derivatives except of the the 0-th one.

The new basis is given as follows. For all \( k, r = 0 \ldots N - 1 \), define

\[
\phi_k^r = e^{i\frac{2\pi}{N} kr} \quad \text{and} \quad \phi_k = (\phi_k^r)_{r=0}^{N-1}.
\tag{2.45}
\]

It is easy to see that the system \( \{\phi_k\}_{k=0}^{N-1} \) is orthonormal with respect to the scalar product \( \langle \cdot, \cdot \rangle \) on \( \mathbb{C}^N \) given by

\[
\langle \zeta, \xi \rangle = \frac{1}{N} \sum_{r=0}^{N-1} \zeta_r^* \xi_r \quad \text{for} \quad \zeta, \xi \in \mathbb{C}^N.
\tag{2.46}
\]

Moreover, for \( x = (x^r)_{r=0}^{N-1} \in \mathbb{R}^N \), let \( \hat{x}(x) = (\hat{x}_k(x))_{k=0}^{N-1} \in \mathbb{C}^N \) be given by

\[
\hat{x}_k(x) = \sqrt{N} \langle x, \phi_k \rangle = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} x^r \phi_k^r.
\tag{2.47}
\]

Note that \( \hat{x}_0(x) = \sqrt{N} P x \) and \( \hat{x}_k(x) = \bar{x}_{N-k}(x) \) for \( k = 1, \ldots, N - 1 \). Hence, the transformation \( x \mapsto \hat{x}(x) \) maps \( \mathbb{R}^N \) into the space

\[
\hat{\mathbb{R}}^N = \mathbb{R} \times \hat{\mathbb{R}}^{N-1},
\tag{2.48}
\]

where

\[
\hat{\mathbb{R}}^{N-1} = \left\{ \hat{x} = (\hat{x}_1, \ldots, \hat{x}_{N-1}) \in \mathbb{C}^{N-1} \mid \hat{x}_k = \bar{x}_{N-k} \text{ for } k = 1, \ldots, N - 1 \right\}.
\tag{2.49}
\]

Furthermore, this map is a bijection. Indeed, for \( \hat{x} = (\hat{x}_k)_{k=0}^{N-1} \in \hat{\mathbb{R}}^N \), the inverse transformation \( x(\hat{x}) = (x^r(\hat{x}))_{r=0}^{N-1} \in \mathbb{R}^N \) is given by

\[
x^r(\hat{x}) = \sqrt{N} \langle \hat{x}, \phi^r \rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k \phi_k^r.
\tag{2.50}
\]

where \( \phi^r = (\phi_k^r)_{k=1}^{N-1} \). It is easy to see that \( x(\hat{x}) \in \mathbb{R}^N \), \( x(\hat{x}(x)) = x \) and \( \sqrt{N} P(x(\hat{x})) = \hat{x}_0 \). Finally, for any \( f \in H^1(\mathbb{R}^N; e^{-H(x)} dx) \), let \( \hat{f} = f \circ x \). Then,

\[
\int_{\mathbb{R}^N} |\nabla f|^2 e^{-H(x)} dx = \int_{\hat{\mathbb{R}}^N} |\nabla \hat{f}|^2 e^{-\sum_{r=0}^{N-1} \frac{1}{r} \psi_j(x^r(\hat{x}))+\frac{1}{2}(\hat{x}_0)^2} d\hat{x}.
\tag{2.51}
\]
We are now in the position to prove Proposition 2.9. Two lemmas that are used in this proof are moved to the Subsections 2.4.1 and 2.4.2.

**Proof of Proposition 2.9.**  
**Step 1.** [Change of basis.]  
Using the Dirichlet principle (Lemma 2.6) and (2.51), we have that

\[
\text{Cap}(B^-, B^+) = \inf_{f : f|_{\partial B^-} = 1, f|_{\partial B^+} = 0} \varepsilon \int_{(B^- \cup B^+)^c} |\nabla f(x)|^2 e^{-H(x)} \, dx \quad (2.52)
\]

\[
= \inf_{f : f|_{\partial B^-} = 1, f|_{\partial B^+} = 0} \varepsilon \int_{\mathbb{R}^N} |\nabla f(\hat{x})|^2 \mathbb{1}_{(\sqrt{N}(-m^*_N + \eta), \sqrt{N}(m^*_N - \eta))}(\hat{x}_0) \, e^{-\sum_{i=0}^{N-1} \frac{1}{2} \psi_J(x^i(\hat{x})) + \frac{1}{2} \frac{J}{\varepsilon} (\hat{x}_0)^2} \, d\hat{x},
\]

where

\[
\hat{B}^- = \{ \hat{x} \in \mathbb{R}^N \mid \hat{x}_0 \leq \sqrt{N}(-m^*_N + \eta) \} \quad \text{and} \quad \hat{B}^+ = \{ \hat{x} \in \mathbb{R}^N \mid \hat{x}_0 \geq \sqrt{N}(m^*_N - \eta) \}. \quad (2.53)
\]

**Step 2.** [Invoking one-dimensional capacities.]  
By the estimate $|\nabla f(\hat{x})|^2 \geq |\partial_0 f(\hat{x})|^2$ and Fubini’s theorem, we infer from (2.52) that

\[
\text{Cap}(B^-, B^+) \geq \inf_{f : f|_{\partial B^-} = 1, f|_{\partial B^+} = 0} \varepsilon \int_{\mathbb{R}^N-1} \sum_{i=0}^{N-1} \psi_J(x^i(\hat{x})) e^{-\frac{1}{2} \sum_{i=0}^{N-1} \frac{1}{2} \psi_J(x^i(\hat{x})) - \frac{1}{2} \frac{J}{\varepsilon} (\hat{x}_0)^2} \, d\hat{x}_0 \, d\hat{x}^\perp. \quad (2.54)
\]

In view of the Dirichlet principle, the quantity inside the integral over $\mathbb{R}^N$ is equal to the one-dimensional capacity of the capacitor $(\sqrt{N}(-m^*_N + \eta), \sqrt{N}(m^*_N - \eta))$. As we have already seen in the proof of Proposition 2.8, we can compute one-dimensional capacities explicitly. Therefore,

\[
\text{Cap}(B^-, B^+) \geq \varepsilon \int_{\mathbb{R}^N-1} \left( \int_{\mathbb{R}^N} \sum_{i=0}^{N-1} \psi_J(x^i(\hat{x})) e^{-\frac{1}{2} \sum_{i=0}^{N-1} \frac{1}{2} \psi_J(x^i(\hat{x})) - \frac{1}{2} \frac{J}{\varepsilon} (\hat{x}_0)^2} \, d\hat{x}_0 \right)^{-1} \, d\hat{x}^\perp. \quad (2.55)
\]

**Step 3.** [Inverse transformation and application of the local Cramér theorem.]  
Using that

\[
\sum_{i=0}^{N-1} \psi_J \left( x^i \left( \left( \sqrt{N} t, \hat{x}^\perp \right) \right) \right) = \sum_{i=0}^{N-1} \psi_J \left( x^i((0, \hat{x}^\perp)) + t \right), \quad (2.56)
\]

\[
= \sum_{i=0}^{N-1} \psi_J \left( x^i((0, \hat{x}^\perp)) \right) + N \psi(t) + N^2 \frac{J}{2} t^2 + \frac{3}{2} \sum_{i=0}^{N-1} x^i((0, \hat{x}^\perp))^2 + t \sum_{i=0}^{N-1} x^i((0, \hat{x}^\perp))^3,
\]

we infer from (2.55) that

\[
\text{Cap}(B^-, B^+) \geq \varepsilon \int_{\mathbb{R}^N-1} d\hat{x}^\perp \, e^{-\sum_{i=0}^{N-1} \frac{1}{2} \psi_J(x^i((0, \hat{x}^\perp)))} \times \left( \int_{\mathbb{R}^N} \psi_J(t) e^{-\frac{3}{2} \sum_{i=0}^{N-1} x^i((0, \hat{x}^\perp))^2 + \frac{3}{2} t \sum_{i=0}^{N-1} x^i((0, \hat{x}^\perp))^3} \, dt \right)^{-1}. \quad (2.57)
\]
We can now use the inverse transformation defined in (2.50) and Proposition 2.1 with $K = [-2, 2]$ to observe that the right-hand side of (2.57) is equal to

$$
\frac{\varepsilon}{\sqrt{N}} \int_{P^{-1}(0)} d\mathcal{H}^{N-1}(x) e^{-\sum_{i=0}^{N-1} \frac{1}{2} \psi_J(x_i)} \left( \int_{-m_i^2 + \eta}^{m_i^2 - \eta} e^{N \frac{1}{2} \psi(t) + \frac{1}{2} t^2 \sum_{i=0}^{N-1} x_i^2 + \frac{1}{2} t \sum_{i=0}^{N-1} x_i^3} dt \right)^{-1} = e^{-NH_x(0)\sqrt{2\pi/N}} \int_{P^{-1}(0)} d\mu_0(x) \left( \int_{-m_i^2 + \eta}^{m_i^2 - \eta} e^{N \frac{1}{2} \psi(t) + \frac{1}{2} t^2 \sum_{i=0}^{N-1} x_i^2 + \frac{1}{2} t \sum_{i=0}^{N-1} x_i^3} dt \right)^{-1} \times \left( 1 + O\left( \frac{1}{\sqrt{N}} \right) \right).
$$

To conclude the proof of this proposition, it remains to show that

$$
\int_{P^{-1}(0)} d\mu_0(x) \left( \int_{-m_i^2 + \eta}^{m_i^2 - \eta} e^{N \frac{1}{2} \psi(t) + \frac{1}{2} t^2 \sum_{i=0}^{N-1} x_i^2 + \frac{1}{2} t \sum_{i=0}^{N-1} x_i^3} dt \right)^{-1} \geq \sqrt{NH_x(0)\sqrt{2\pi/N}} \left( 1 + O\left( \sqrt{\frac{\log(N)}{N}} \right) + O(\varepsilon^2) \right).
$$

**Step 4.** [Proof of (2.59).]

We show (2.59) in three steps.

**Step 4.1.** [Some notation.]

Recall the definition of $\mu^{\varepsilon,0}$ (see (2.10)). Then, let

$$
M_k = \int_R z^k d\mu^{\varepsilon,0}(z) \quad \text{for } k \in \mathbb{N}, \quad \text{and}
$$

$$
\Gamma = \left\{ x \in P^{-1}(0) \bigg| \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \leq \frac{1}{\varepsilon} M_2 N + \frac{1}{\varepsilon} \sqrt{M_4 - (M_2)^2} \sqrt{N \log(N)}, \right. \ \left. \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^3 \right\} \leq \frac{1}{\varepsilon} \sqrt{M_6 - (M_4)^2} \sqrt{N \log(N)}
$$

Note that, by applying Jensen’s inequality to the function $(r, s) \mapsto r^2/s$, we have that $M_6 - (M_4)^2/M_2 > 0$.

Let $\text{lhs}[(2.59)]$ denote the left-hand side of (2.59). Then,

$$
\text{lhs}[(2.59)] \geq \int_{\Gamma} d\mu_0(x) \left( \int_{-m_i^2 + \eta}^{m_i^2 - \eta} e^{N \frac{1}{2} \psi(t) + \frac{1}{2} t^2 \sum_{i=0}^{N-1} x_i^2 + \frac{1}{2} t \sum_{i=0}^{N-1} x_i^3} dt \right)^{-1} = \int_{\Gamma} d\mu_0(x) \left( \int_{-m_i^2 + \eta}^{m_i^2 - \eta} e^{\sum_{i=0}^{N-1} F_x(t)} dt \right)^{-1},
$$

where for $x \in P^{-1}(0)$ and $m \in K$,

$$
F_x(t) = \frac{1}{4} t^4 - \left( 1 - 3M_2 - \frac{\sqrt{\log(N)}}{\sqrt{N}} \frac{3}{2} \sqrt{M_4 - (M_2)^2} \right) \frac{1}{2} t^2 + t \frac{1}{N} \sum_{i=0}^{N-1} x_i^3 = \frac{1}{4} t^4 - (1 + \alpha_{N, \varepsilon}) \frac{1}{2} t^2 + t q(x).
$$
Using that, by Corollary B.2, \( M_k = \Omega(\varepsilon^{\frac{k}{2}}) \) for \( k = 2, 4, 6 \), we observe that \( \alpha_{N,\varepsilon} = O(\varepsilon) \) and, for \( x \in \Gamma \), \( q(x) = \Omega(\varepsilon^2 \sqrt{\log(N) / N}) \). Therefore, by applying Cardano’s method (see [9, p. 22]), we have that

- \( F_x \) admits a unique local maximum at \( z_1 = q(x)/(1 + \alpha_{N,\varepsilon}) + O(q(x)^2) = O\left( \varepsilon^2 \sqrt{\log(N) / N} \right) \), and

- \( F_x \) admits exactly two local minima at some points \( z_2 = 1 + O\left( \varepsilon^2 \sqrt{\log(N) / N} \right) \) and \( z_3 = -1 + O\left( \varepsilon^2 \sqrt{\log(N) / N} \right) \).

**Step 4.2.** [Restriction of \([-m_\ast + \eta, m_\ast - \eta]\) to a small neighbourhood of \( z_1 \).]

In the following, \( C \) denotes a varying positive constant, which is independent of \( \varepsilon \) and \( N \). Let

\[
\rho = \sqrt{\frac{3 \varepsilon \log(N\varepsilon^{-1})}{N|F_x' (z_1)|}}.
\]

Then, by Taylor’s formula, for some \( \theta \in [0, 1] \) and by the estimates on \( q(x) \) for \( x \in \Gamma \),

\[
e^\frac{N}{\pi} F_x (z_1 + \rho) = e^{\frac{N}{\pi} F_x (z_1)} e^{\frac{N}{2\pi} F_x'' (z_1)} e^{\frac{N}{6\pi} F_x''' (z_1 + \theta \rho)}
\]

\[
e = e^{\frac{N}{3\pi} \left( \frac{q(x)^2}{N^{1+\alpha_{N,\varepsilon}}} + O(q(x)^3) \right)} N^{-\frac{3}{2}} e^{\frac{N}{3\pi} \rho (z_1 + \theta \rho)} \leq C N^{\frac{3}{2}} N^{-\frac{3}{2}} e^{\frac{3}{2}} \leq C N^{-1} \varepsilon^3.
\]

Similarly, we can show that \( e^{\frac{N}{\pi} F_x (z_1 - \rho)} \leq C N^{-1} \varepsilon^2 \). Then, since \( \sup_{m \in [-m_\ast + \eta, z_1 - \rho]} F_x(t) = F_x(z_1 - \rho) \) and \( \sup_{m \in [z_1 + \rho, m_\ast - \eta]} F_x(t) = F_x(z_1 + \rho) \),

\[
\int_{[-m_\ast + \eta, m_\ast - \eta] \setminus [z_1 - \rho, z_1 + \rho]} e^{\frac{N}{\pi} F_x(t)} dt \leq m_\ast \left( e^{\frac{N}{\pi} F_x (z_1 + \rho)} + e^{\frac{N}{\pi} F_x (z_1 - \rho)} \right) \leq C N^{-1} \varepsilon^2.
\]

By using Jensen’s inequality, this implies that

\[
\text{lhs}[2.59] \geq \int_{\Gamma} d\mu_0(x) \left( \int_{z_1 - \rho}^{z_1 + \rho} e^{\frac{N}{\pi} F_x(t)} dt + C N^{-1} \varepsilon^2 \right)^{-1} \geq \left( \frac{1}{\mu_0(\Gamma)} \int_{\Gamma} d\mu_0(x) \int_{z_1 - \rho}^{z_1 + \rho} e^{\frac{N}{\pi} F_x(t)} dt + C N^{-1} \varepsilon^2 \right)^{-1} \mu_0(\Gamma).
\]

**Step 4.3.** [Conclusion of the proof of (2.59).]

Note that \( \mu_0(\Gamma) = 1 + O(1/\sqrt{N}) \) (see Lemma 2.12) and \( |\tilde{h}_x''(0)| = O(1/\varepsilon) \) (see Step 3 of the proof of Proposition 2.8). Therefore, (2.66) implies (2.59) once we show that

\[
\int_{\Gamma} d\mu_0(x) \int_{z_1 - \rho}^{z_1 + \rho} e^{\frac{N}{\pi} F_x(t)} dt \leq \frac{\sqrt{2\pi}}{\sqrt{N|H_x''(0)|}} \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) + O(\varepsilon^2) \right).
\]

Since \([2\rho, -2\rho] \supset [z_1 - \rho, z_1 + \rho] \) for \( \varepsilon \) small enough, we have that

\[
\int_{\Gamma} d\mu_0(x) \int_{z_1 - \rho}^{z_1 + \rho} e^{\frac{N}{\pi} F_x(t)} dt \leq \int_{-2\rho}^{2\rho} e^{N^{\frac{1}{2}} \left( 1 + \alpha_{N,\varepsilon} \right) \varepsilon^2 \phi_{N,\varepsilon}(t, 0)} dt e^{N\phi_{N,\varepsilon}(0)},
\]

18
where \( \varphi_{N,\epsilon} \) was defined in (2.6), and \( \phi_{N,\epsilon} : \mathbb{R}^2 \to \mathbb{R} \) is given by

\[
\phi_{N,\epsilon}(t, m) = -\frac{1}{N} \log \int_{P^{-1}(m)} e^{-\frac{1}{\epsilon} \sum_{i=0}^{N-1}(\psi_J(x_i) - mx_i^2)} \mathcal{H}^{N-1}(dx). \quad (2.69)
\]

For all \( t \in [-2\rho, 2\rho] \), the function \( z \mapsto \psi_J(z) - t^3 \) fulfills the conditions of Proposition C.2.

Therefore, by Proposition 2.1 and Proposition C.2,

\[
\int \mu_0(x) \int_{z_1}^{z_{1+\rho}} e^{\frac{N}{\epsilon} F_y(t)} dt \leq \int_{-2\rho}^{2\rho} e^{\frac{N}{\epsilon} \left[ \frac{t^4}{12} - \frac{1}{2} \phi(t, 0) \right]} \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right),
\]

(2.70)

where \( \phi(t, \cdot) \) is the Legendre transform of the function

\[
\sigma \in \mathbb{R} \mapsto \phi^*(t, \sigma) = \log \int_{\mathbb{R}} e^{\sigma z + \frac{1}{12} t z^3 - \frac{1}{2} \psi_J(z)} dz.
\]

(2.71)

In Lemma 2.11 below we show that for all \( t \in [-2\rho, 2\rho] \),

\[
\phi(t, 0) = \varphi(0) - \frac{t^2}{2\epsilon^2} \left( M_6 - \frac{(M_4)^2}{M_2} \right) + \frac{t^3}{6} O_K(\epsilon^2), \quad \text{and}
\]

\[
\partial_{t^2} \phi(t, 0) = \varphi''(0) \left( 1 + O_K \left( \epsilon \sqrt{\log(N) / N} \right) \right).
\]

(2.72)

Combining (2.72) with (2.70), and using that \( t^3, t^4 \in O_K(\epsilon N^{-\frac{1}{2}} \sqrt{\log(N)^3}) \) for \( t \in [-2\rho, 2\rho] \), we have that

\[
\int \mu_0(x) \int_{z_1}^{z_{1+\rho}} e^{\frac{N}{\epsilon} F_y(t)} dt \leq \int_{-2\rho}^{2\rho} e^{\frac{N}{\epsilon} \left[ \frac{t^4}{12} - \frac{1}{2} \phi(t, 0) \right]} \left( 1 + O \left( \frac{\log(N)^3}{N} \right) \right) \quad (2.73)
\]

To conclude (2.67), it only remains to show that

\[
\frac{1}{\epsilon} \left( 1 + \alpha_{N,\epsilon} - \frac{1}{\epsilon} \left( M_6 - \frac{(M_4)^2}{M_2} \right) \right) = \left( 1 + O \left( \epsilon \sqrt{\log(N) / N} \right) + O(\epsilon^2) \right) |H''_{\epsilon}(0)|.
\]

(2.74)

From a simple computation via integration by parts, we have that

\[
\frac{1}{\epsilon} \left( 1 - 3M_2 + \frac{1}{\epsilon} \left( M_6 - \frac{(M_4)^2}{M_2} \right) \right) = \frac{1}{\epsilon} J - (M_2)^{-1} = |H''_{\epsilon}(0)|.
\]

(2.75)

Moreover, using the fact that \( M_k = \Omega(\epsilon^k) \) for \( k = 2, 4, 6 \), we have that

\[
\alpha_{N,\epsilon} = -3M_2 + O \left( \epsilon \sqrt{\log(N) / N} \right) \quad \text{and} \quad \frac{1}{\epsilon} \left( M_6 - \frac{(M_4)^2}{M_2} \right) = \Omega(\epsilon^2).
\]

(2.76)

Combining (2.75) and (2.76) implies (2.74). This concludes the proof of this proposition. \( \square \)
Remark 2.10  In the proof of Proposition 2.9, we have used that, under $\mu_0$, the fluctuations of $\frac{1}{N} \sum_i x_i^3$ can be estimated by

$$R(\varepsilon, N) := \sqrt{\frac{\log(N)}{N}} \sqrt{\frac{M_6 - (M_4)^2}{M_2}}.$$  

(2.77)

This is the content of Lemma 2.12, which can be shown for all $J > 0$. In the case $J > 1$, we have that $R(\varepsilon, N) = \Omega(\varepsilon^3 \sqrt{\log(N)/N})$. Hence, under $\mu_0$, $\frac{1}{N} \sum_i x_i^3$ is concentrated around 0 for large $N$ and for small $\varepsilon$. Heuristically, this can be interpreted by saying that we can control the fluctuations at the macroscopic saddle point 0 in the regime $J > 1$, $N \to \infty$ and $\varepsilon \ll 1$. However, in the case $J < 1$, we have that $R(\varepsilon, N) = \Omega(\sqrt{\log(N)/N})$. Therefore, we are not able to control the fluctuations of $\frac{1}{N} \sum_i x_i^3$ for small $\varepsilon$, and the methods of this proof can not be used to prove Proposition 2.9 for $J < 1$. This is the main reason why we made the assumption that $J > 1$ (i.e. (1.8)).

2.4.1 Taylor expansion of $\phi_\varepsilon$

Lemma 2.11 Using the same notation as in the proof of Proposition 2.9, we have that for all $t \in [-2\rho, 2\rho]$,

$$\phi_\varepsilon(t, 0) = \varphi_\varepsilon(0) - \frac{t^2}{2\varepsilon^2} \left( M_6 - \frac{(M_4)^2}{M_2} \right) + \frac{t^3}{6} O_K(\varepsilon^2), \quad \text{and}$$

$$\partial_{22} \phi_\varepsilon(t, 0) = \varphi''_\varepsilon(0) \left( 1 + O_K \left( \varepsilon \sqrt{\frac{\log(N)}{N}} \right) \right).$$  

(2.78)

Proof. The computations are quite involved, so it is useful to simplify the notation. We write

$$\phi(t, \sigma) = \phi_\varepsilon(t, \sigma) \quad \text{and} \quad \phi^*(t, \sigma) = \phi_\varepsilon^*(t, \sigma),$$  

(2.80)

and for all $t \in K := [-2, 2]$ and for all $k \in \mathbb{N}$, let

$$M_{k, t} = \int_{\mathbb{R}} e^{2\phi(t, 0) x + \varepsilon x^3 - \frac{1}{2} \psi_J(x)} dx \quad \text{and} \quad M_k := M_{k, 0}.$$  

(2.81)

Then, (A.4) can be written as

$$\phi(t, \sigma) = \partial_2 \phi(t, \sigma) \sigma - \phi^*(t, \partial_2 \phi(t, \sigma)).$$  

(2.82)

The proofs of (2.78) and (2.79) follow from Taylor’s formula and are given in seven steps. In the first four steps, we compute $\partial_1 \phi(t, 0), \partial_{11} \phi(t, 0)$ and $\partial_{111} \phi(t, 0)$. In the fifth step, we show that the remainder term in Taylor’s formula is negligible. In the sixth and seventh step we show (2.78) and (2.79), respectively.

Step 1. [Computation of the partial derivatives of $\phi^*$.] In order to compute $\partial_1 \phi(t, 0), \partial_{11} \phi(t, 0)$ and $\partial_{111} \phi(t, 0)$, we need to compute several partial
derivatives of $\phi^*$ at the point $(t, \partial_2 \phi(t, 0))$ (see (2.92)). The formulas follow from straightforward computations, and are given as follows.

$$
\begin{align*}
\partial_1 \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon} M_{3, t}, \\
\partial_{11} \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon} \left( M_{6, t} - (M_{3, t})^2 \right), \\
\partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon} M_{4, t}, \\
\partial_{111} \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon^2} (M_{9, t} - 3 M_{6, t} M_{3, t} + 2 (M_{3, t})^3), \\
\partial_{112} \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon^2} (M_{7, t} - 2 M_{4, t} M_{3, t}), \\
\partial_{122} \phi^*(t, \partial_2 \phi(t, 0)) &= \frac{1}{\varepsilon} (M_{5, t} - M_{2, t} M_{3, t}).
\end{align*}
$$

(2.83)

Here we have used several times that $\int_{\mathbb{R}} e^{\partial_2 \phi(t, 0) z + \frac{1}{2} t^3 - \frac{1}{2} \psi_2(z)} \, dz = 0$.

**Step 2.** [Computation of $\partial_1 \phi$.]

Using (2.82), we compute that

$$
\partial_1 \phi(t, \sigma) = - \partial_1 \phi^*(t, \partial_2 \phi(t, \sigma)).
$$

(2.84)

Here we have used that, by (A.4), $\partial_2 \phi^*(t, \partial_2 \phi(t, \sigma)) = \sigma$. In particular,

$$
\partial_1 \phi(t, 0) = - \partial_1 \phi^*(t, \partial_2 \phi(t, 0)) \quad \text{and} \quad \partial_1 \phi(0, 0) = 0.
$$

(2.85)

**Step 3.** [Computation of $\partial_{11} \phi(t, \sigma)$ and $\partial_{11}(0, 0)$.

Taking the derivative with respect to $t$ in (2.85), we obtain that

$$
\partial_{11} \phi(t, 0) = - \partial_{11} \phi^*(t, \partial_2 \phi(t, 0)) - \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{12} \phi(t, 0).
$$

(2.86)

And in order to compute $\partial_{12} \phi(t, 0)$, we take the derivative with respect to $\sigma$ in (2.84) and use (2.83) and Lemma A.1 to observe that

$$
\partial_{12} \phi(t, 0) = - \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{22} \phi(t, 0) = - \frac{1}{\varepsilon} M_{4, t} (M_{2, t})^{-1}.
$$

(2.87)

Therefore,

$$
\partial_{11} \phi(t, 0) = - \partial_{11} \phi^*(t, \partial_2 \phi(t, 0)) + (\partial_{12} \phi^*(t, \partial_2 \phi(t, 0)))^2 \partial_{22} \phi(t, 0).
$$

(2.88)

and for $m = 0$,

$$
\partial_{11} \phi(0, 0) = - \frac{1}{\varepsilon^2} (M_6 - (M_3)^2 (M_2)^{-1}).
$$

(2.89)

**Step 4.** [Computation of $\partial_{111} \phi(t, 0).$]

Taking the derivative with respect to $t$ in (2.88) and using (2.87), we obtain that

$$
\begin{align*}
\partial_{111} \phi(t, 0) &= - \partial_{111} \phi^*(t, \partial_2 \phi(t, 0)) - \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{12} \phi(t, 0) \\
&\quad + (\partial_{12} \phi^*(t, \partial_2 \phi(t, 0)))^2 \partial_{22} \phi(t, 0) \\
&\quad + 2 \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{12} \phi(t, 0) \partial_{22} \phi(t, 0) \\
&\quad + 2 \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{11} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{22} \phi(t, 0) \\
&\quad = - \partial_{111} \phi^*(t, \partial_2 \phi(t, 0)) + (\partial_{12} \phi^*(t, \partial_2 \phi(t, 0)))^2 \partial_{22} \phi(t, 0) \\
&\quad - 2 (\partial_{12} \phi^*(t, \partial_2 \phi(t, 0)))^2 \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) (\partial_{22} \phi(t, 0))^2 \\
&\quad + 3 \partial_{12} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{11} \phi^*(t, \partial_2 \phi(t, 0)) \partial_{22} \phi(t, 0).
\end{align*}
$$

(2.90)
We see that we also have to compute \( \partial_{122}(t, 0) \). Using the relations (A.5) and (2.87), we have that

\[
\partial_{122}(t, 0) = \frac{d}{dt} (\partial_{22}(t, 0))^{-1} = \partial_{22}(t, 0) \partial_{12}(t, 0) (\partial_{22}(t, 0))^3
\]

(2.91)

Therefore, by plugging (2.91) into (2.90) and by using Lemma A.1 and (2.83), we have that

\[
\partial_{111}(t, 0) = -\partial_{111}(t, 0) + \partial_{22}(t, 0) (\partial_{12}(t, 0))^2 (\partial_{22}(t, 0))^3
\]

(2.92)

\[
\bigg(\int_{\mathbb{R}} \left( M_{9,t} - 3M_{6,t}3t + 2(M_{3,t})^3 - M_{3,t}(M_{4,t})^3(M_{2,t})^{-3} \right)
\]

\[
- \frac{3}{\varepsilon^3} (M_{5,t} - 2M_{2,t}3M_{3,t}) (M_{4,t})^2 (M_{2,t})^{-2}
\]

\[
+ \frac{3}{\varepsilon^3} (M_{7,t} - 2M_{4,t}3M_{3,t}) M_{4,t} (M_{2,t})^{-1}.
\]

**Step 5.** [Estimation of \( \partial_{111}(t, 0) \).]

Note that for all \( t \in [-2\rho, 2\rho] \), the function \( z \mapsto \psi_J(z) - tz^3 \) fulfils the conditions of Lemma C.1. Hence, there exist \( \varepsilon_0 > 0 \) and a function \( \tau_{\varepsilon,t} \) such that

\[
sup_{\varepsilon \in (0, \varepsilon_0)} \sup_{t \in [-2\rho, 2\rho]} |\tau_{\varepsilon, t}(0)| < \infty \quad \text{and} \quad \partial_{2\phi}(t, 0) = \frac{1}{\varepsilon} \tau_{\varepsilon, t}(0),
\]

(2.93)

Moreover, for all \( t \in [-2\rho, 2\rho] \), the function \( U(t, z) := \tau_{\varepsilon,t}(0)z + \psi_J(z) - tz^3 \) fulfils the conditions of Corollary B.2. The dependence on \( \varepsilon \) of \( \tau_{\varepsilon,t}(0) \) is not a problem here due to its uniform boundedness in \( \varepsilon \) stated in (2.93). Therefore, by Corollary B.2, for all \( k \in \mathbb{N} \),

\[
M_{k,t} = \frac{\int_{\mathbb{R}} \left( z - \frac{\int_{\mathbb{R}} e^{2\phi(t,0)r + \frac{1}{2}tr^2 - \frac{1}{2}t^3 - \frac{1}{2}t^3} \psi_J(r) dr}{\int e^{2\phi(t,0)r + \frac{1}{2}tr^2 - \frac{1}{2}t^3} \psi_J(r) dr} \right)^k e^{2\phi(t,0)z + \frac{1}{2}t^3 - \frac{1}{2}t^3} \psi_J(z) dz}{\int e^{2\phi(t,0)r + \frac{1}{2}tr^2 - \frac{1}{2}t^3} \psi_J(r) dr} = \Omega_K(\varepsilon^{\frac{1}{2}}).
\]

(2.94)

Hence, in view of (2.92),

\[
\partial_{111}(t, 0) = O_K(\varepsilon^2).
\]

(2.95)

**Step 6.** [Conclusion of the proof of (2.78).]

By Taylor’s formula, for some \( \theta \in [0, 1] \), and by (2.85), (2.89) and (2.95),

\[
\phi(t, 0) = \phi(0, 0) + m \partial_1\phi(0, 0) + \frac{t^2}{2} \partial_{11}\phi(0, 0) + \frac{t^3}{6} \partial_{111}\phi(\theta m, 0)
\]

\[
= \varphi_t(0) - \frac{t^2}{2\varepsilon^2} \left( M_6 - \frac{(M_4)^2}{M_2} \right) + t^3 O_K(\varepsilon^2).
\]

(2.96)
Step 7. [Conclusion of the proof of (2.79).] Again, by Taylor’s formula, for some \( \theta \in [0, 1] \),

\[
\partial_{22}\phi(t, 0) = \varphi''(0) \left( 1 + t \frac{\partial_{122}\phi(\theta t, 0)}{\varphi''(0)} \right).
\]

(2.97)

Using (2.83), (2.91), (2.94) and (2.38), we have that

\[
\left| t \frac{\partial_{122}\phi(\theta t, 0)}{\varphi''(0)} \right| \leq \epsilon |t| |M_{5,\theta t} M_{4,\theta t} (M_{2, \theta t})^{-3} - (M_{5,\theta t} - M_{2,\theta t} M_{3, \theta t}) (M_{2, \theta t})^{-2}|
\leq C \rho \varepsilon
\]

for some \( C > 0 \), independent of \( \varepsilon \) and \( t \in [-2\rho, 2\rho] \). This concludes the claim. \( \square \)

2.4.2 LDP estimates for the conditional measure at the saddle

Lemma 2.12 Using the same notation as in the proof of Proposition 2.9, we have that

\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \geq \frac{M_2}{\varepsilon} N + \frac{1}{\varepsilon} \sqrt{M_4 - (M_2)^2} \sqrt{N \log(N)} \right) = O \left( \frac{1}{\sqrt{N}} \right), \quad \text{and}
\]

(2.99)

\[
\mu_0 \left( \frac{1}{\varepsilon} \left| \sum_{k=0}^{N-1} x_k^3 \right| \geq \frac{1}{\varepsilon} \sqrt{M_6 - (M_2)^2} \sqrt{N \log(N)} \right) = O \left( \frac{1}{\sqrt{N}} \right).
\]

(2.100)

Remark 2.13 The same arguments that we use in the proof of Lemma 2.12 also show that

\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \leq \frac{M_2}{\varepsilon} N - \frac{1}{\varepsilon} \sqrt{M_4 - (M_2)^2} \sqrt{N \log(N)} \right) = O \left( \frac{1}{\sqrt{N}} \right).
\]

(2.101)

Proof of Lemma 2.12. Let \( t \in (0, T_0) \) for some \( T_0 \in (0, 2) \). In this proof, \( C \) denotes a varying positive constant, which does not depend on \( \varepsilon, N \) and \( t \) (but may depend on \( T_0 \)).

We use similar arguments here as in the proofs of Proposition 2.9 and Lemma 2.11. We first show \((2.99)\). Let \( C_{\varepsilon, N} = M_2/\varepsilon + \alpha_\varepsilon \sqrt{\log(N)/\sqrt{N}} \), where \( \alpha_\varepsilon = \sqrt{M_4 - (M_2)^2}/\varepsilon \). Using the exponential Markov inequality, we infer that

\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \geq C_{\varepsilon, N} \right) \leq \int_{P_{-1}(0)} e^{\frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2} d\mu_0(x) = \frac{e^{N(\varphi_{N,\varepsilon}(0) - \omega_{N,\varepsilon}(t,0))}}{e^{NC_{N,\varepsilon}}},
\]

(2.102)

where \( \varphi_{N,\varepsilon} \) was defined in \((2.6)\), and \( \omega_{N,\varepsilon} : \mathbb{R}^2 \to \mathbb{R} \) is given by

\[
\omega_{N,\varepsilon}(t, m) = -\frac{1}{N} \log \int_{P_{-1}(m)} e^{-\frac{1}{\varepsilon} \sum_{k=0}^{N-1} (\frac{1}{2} \psi_{J}(x_k) - tx_k^2)} \mathcal{H}^N(dx).
\]

(2.103)

If \( T_0 \in (0, 2) \) is small enough, the function \( z \mapsto V_z(t) := \frac{1}{\varepsilon} \psi_{J}(z) - tz^2 \) satisfies the assumptions of Proposition C.2. Hence, by applying Proposition 2.1 and Proposition C.2,

\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \geq C_{\varepsilon, N} \right) \leq C e^{N(\varphi_{N}(0) - \omega_{N}(t,0)) - tC_{N, \varepsilon}} \sqrt{\frac{\partial_{22}\omega_{\varepsilon}(t, 0)}{\varphi''(0)}},
\]

(2.104)
where $\omega_\varepsilon(t, \cdot)$ is the Legendre transform of the function
\[
\sigma \in \mathbb{R} \mapsto \omega_\varepsilon^*(t, \sigma) = \log \int_{\mathbb{R}} e^{\sigma z + \frac{1}{2} t z^2 - \frac{1}{4} \psi_j(z)} \, dz. \tag{2.105}
\]
Using (A.5), Lemma C.1 (iii) and the fact that $|t| < T_0$, we have that
\[
\sup_{t \in (0, T_0)} \sqrt{\frac{\partial_{22} \omega_\varepsilon(t, 0)}{\phi_\varepsilon''(0)}} \leq C. \tag{2.106}
\]
Moreover, by Taylor’s formula, for some $\theta \in [0, 1]$, \[
\omega_\varepsilon(t, 0) = \varphi_\varepsilon(0) + t \partial_1 \omega_\varepsilon(0, 0) + \frac{t^2}{2} \partial_{11} \omega_\varepsilon(0, 0) + \frac{t^3}{6} \partial_{111} \omega_\varepsilon(\theta t, 0). \tag{2.107}
\]
Since $V$ is even, we have that $\partial_2 \omega_\varepsilon(0, 0) = 0$. Therefore, in view of (A.4), we have that
\[
\omega_\varepsilon(t, 0) = -\omega_\varepsilon^*(t, 0) = -\log \int_{\mathbb{R}} e^{\frac{1}{2} t z^2 - \frac{1}{4} \psi_j(z)} \, dz, \tag{2.108}
\]
and a straightforward computation yields that
\[
\partial_1 \omega_\varepsilon(0, 0) = -\frac{1}{\varepsilon} M_2, \quad \partial_{11} \omega_\varepsilon(0, 0) = -\frac{1}{\varepsilon^2} M_4 + \frac{1}{\varepsilon^2} (M_2)^2 = -\alpha_\varepsilon^2, \quad \text{and}
\]
\[
\partial_{111} \omega_\varepsilon(\theta t, 0) = -\frac{1}{\varepsilon^3} \int_{\mathbb{R}} z^6 e^{\frac{1}{2} \alpha t z^2 - \frac{1}{4} \psi_j(z)} \, dz - \frac{2}{\varepsilon^3} \left( \int_{\mathbb{R}} e^{\frac{1}{2} \alpha t z^2 - \frac{1}{4} \psi_j(z)} \, dz \right)^3 \tag{2.109}
\]
\[
+ \frac{3}{\varepsilon^3} \int_{\mathbb{R}} e^{\frac{1}{2} \alpha t z^2 - \frac{1}{4} \psi_j(z)} \, dz \int_{\mathbb{R}} e^{\frac{1}{2} \alpha t z^2 - \frac{1}{4} \psi_j(z)} \, dz \int_{\mathbb{R}} e^{\frac{1}{2} \alpha t z^2 - \frac{1}{4} \psi_j(z)} \, dz.
\]
Using Corollary B.2, we observe that $\sup_{t \in (0, T_0)} |\partial_{111} \omega_\varepsilon(\theta t, 0)| \leq C$. Hence, combining all the previous estimates,
\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^2 \geq C_{\varepsilon, N} \right) \leq C e^{N(t \frac{1}{2} M_2 + \frac{\varepsilon^2}{2} \alpha_\varepsilon^2 + \frac{\varepsilon^3}{6} C - t C_{\varepsilon, N})} = C e^{N(\frac{2}{3} \alpha_\varepsilon^2 - t \sqrt{\log(N)} \alpha_\varepsilon)}. \tag{2.110}
\]
Choosing $t = \sqrt{\log(N)} \sqrt{N}^{-1} \alpha_\varepsilon^{-1}$ concludes the proof of (2.99).

To show (2.100), we first use the same arguments as in the proof of (2.99), and infer that
\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^3 \geq \frac{1}{\varepsilon} \sqrt{M_6 - \frac{(M_4)^2}{M_2} \sqrt{N \log(N)}} \right) = O \left( \frac{1}{\sqrt{N}} \right). \tag{2.111}
\]
Indeed, here, in contrast to the proof of (2.99), we have to use the functions $\phi_{N, \varepsilon}$ and $\phi_\varepsilon$ (recall their definitions in the proof of Proposition 2.9) instead of $\omega_{N, \varepsilon}$ and $\omega_\varepsilon$, and we use (2.78) instead of (2.109).

It remains to show the other inequality in (2.100). This is also shown in the same way as before. Let $c_{\varepsilon, N} = \sqrt{M_6 - (M_4)^2 / M_2} \sqrt{\log(N) / (\varepsilon \sqrt{N})}$. First, we use the exponential Markov inequality to observe that for $t > 0$,
\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^3 \leq -c_{\varepsilon, N} N \right) \leq e^{N(\varphi_{N, \varepsilon}(0) - \phi_{N, \varepsilon}(-t, 0) - t c_{\varepsilon, N})}. \tag{2.112}
\]
Then, as before, we apply Proposition C.2 and (2.78), we show (the analogue of) (2.106), and choose \( t = \sqrt{\log(N)} / (\sqrt{M_0 - (M_4)^2/M_2} \sqrt{N} \varepsilon) \) to show that

\[
\mu_0 \left( \frac{1}{\varepsilon} \sum_{k=0}^{N-1} x_k^3 \leq -c\epsilon N \right) \leq \frac{C}{\sqrt{N}},
\]

(2.113)

We conclude the proof of this lemma. \( \square \)

2.5 The mass of the equilibrium potential

**Proposition 2.14** Consider the same setting as in Theorem 2.7. Then, for \( \varepsilon \) small enough,

\[
\int_{(B^+)^c} f_{B^-,B^+}^e(y) e^{-H(y)} dy = e^{-N H_\epsilon(-m^*_\varepsilon)} \frac{\sqrt{\varphi''_\epsilon(-m^*_\varepsilon)}}{\sqrt{H''_\epsilon(-m^*_\varepsilon)}} \left( 1 + O \left( \sqrt{\frac{\log(N)^3}{N}} \right) \right).
\]

(2.114)

**Proof.** In this proof, \( C \) denotes a varying positive constant, which is independent of \( \varepsilon \) and \( N \).

**Step 1.** [Splitting into four regions.]

Let \( R > 2 \) be a positive number, which is independent of \( N, \varepsilon \) and \( m \in K \), and whose precise value is chosen later in Step 3. Using that \( f_{B^-,B^+}^e(y) = 1 \) for \( y \in B^- \), we split the left-hand side of (2.114) according to this \( R \) in the following way.

\[
\int_{(B^+)^c} f_{B^-,B^+}^e(y) e^{-H(y)} dy = \int_{\{P \in [-m^*_\varepsilon + \eta,-m^*_\varepsilon + \eta]\}} e^{-H(y)} dy + \int_{\{P \in [-m^*_\varepsilon + \eta,m^*_\varepsilon - \eta]\}} f_{B^-,B^+}^e(y) e^{-H(y)} dy + \int_{\{P \in [-R^-,m^*_\varepsilon + \eta]\}} e^{-H(y)} dy + \int_{\{x \in \mathbb{R}^N \mid P_x < -R\}} e^{-H(y)} dy =: I + II + III + IV.
\]

(2.115)

In Step 2 we compute the asymptotic value of the term \( I \), and in Step 3 and 4 we show that the terms \( II, III \) and \( IV \) are of lower order than \( I \).

**Step 2.** [Estimation of the term \( I \).]

Note that, using the same arguments as in Step 4 and Step 5 of the proof of Proposition 2.8, for all \( m \in [-m^*_\varepsilon - \eta,-m^*_\varepsilon + \eta] \), (recall the definition of \( \eta \) in (2.26)),

\[
\sqrt{\varphi''_\epsilon(m)} = \sqrt{\varphi''_\epsilon(-m^*_\varepsilon)} \left( 1 + O_K \left( \sqrt{\frac{\log(N\varepsilon^{-1})}{N}} \right) \right), \quad \text{and}
\]

\[
H_\epsilon(m) = H_\epsilon(-m^*_\varepsilon) + \frac{1}{2} (m + m^*_\varepsilon)^2 H''_\epsilon(-m^*_\varepsilon) + O_K \left( \sqrt{\varepsilon} \sqrt{\frac{\log(N\varepsilon^{-1})}{N}} \right).
\]

(2.116)

Then, using the coarea formula in the same way that we did in Subsection 2.1.1 and applying Proposition 2.1 for the compact set \( L = [-R,R] \), we observe that

\[
I = \sqrt{N} \int_{-m^*_\varepsilon - \eta}^{-m^*_\varepsilon + \eta} e^{-N \varphi_N,\epsilon(m)} \frac{1}{\sqrt{2\pi}} \frac{N m^2}{2} dm = \sqrt{N} \int_{-m^*_\varepsilon - \eta}^{-m^*_\varepsilon + \eta} e^{-N H_\epsilon(m)} \frac{\sqrt{\varphi''_\epsilon(m)}}{\sqrt{2\pi}} dm \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right).
\]

(2.117)
Using (2.116) and arguing as in the proof of Proposition 2.8, we have that for \( \varepsilon \) small enough,

\[
I = \frac{\sqrt{N}}{\sqrt{2\pi}} e^{-NR_e(-m_\varepsilon^2)} \sqrt{\varphi'_\varepsilon(-m_\varepsilon^2)} \int_{-\eta}^{\eta} e^{-NR''_e(-m_\varepsilon^2)m_\varepsilon^2} \, dm \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) \right)
\]

\[
= e^{-NR_e(-m_\varepsilon^2)} \sqrt{\varphi'_\varepsilon(-m_\varepsilon^2)} \left( 1 + O\left( \sqrt{\frac{\log(N)^3}{N}} \right) \right). \tag{2.118}
\]

**Step 2.** [Estimation of the terms II and III.] We only consider the term II. The term III can be estimated in the same way. By using that \( |f_{B^-B^+}^*| \leq 1 \) and by applying the coarea formula and Proposition 2.1 as in Step 1, we have that

\[
|II| \leq C \sqrt{N} \int_{-m_\varepsilon^2+\eta}^{m_\varepsilon^2-\eta} e^{-NR_e(m)} \sqrt{\varphi'_\varepsilon(m)} \, dm. \tag{2.119}
\]

Note that, similarly as in Step 3 of the proof of Proposition 2.8, we have that \( |\tilde{H}'_\varepsilon(-m_\varepsilon^2)| = \Omega(1/\varepsilon) \). Together with (2.38), this shows that \( I = \Omega(e^{-NR_e(-m_\varepsilon^2)}) \). In the following we prove that \( II \leq O(e^{-NR_e(-m_\varepsilon^2)}\sqrt{N}^{-1}) \), which shows that II is of lower order than I. Since \( \tilde{H}_\varepsilon \) is symmetric and has its two global minima at \( \pm m_\varepsilon^* \), we have that

\[
\inf_{m \in [-m_\varepsilon^*+\eta, m_\varepsilon^*-\eta]} \tilde{H}_\varepsilon(m) = \tilde{H}_\varepsilon(-m_\varepsilon^*+\eta). \tag{2.120}
\]

Then, by (2.38), (2.116) and the definition of \( \eta \) (see (2.26)),

\[
|I_2| \leq \frac{C \sqrt{N}}{\sqrt{\varepsilon}} e^{-NR_e(-m_\varepsilon^2+\eta)} \leq \frac{C \sqrt{N}}{\sqrt{\varepsilon}} e^{-N(R_e(-m_\varepsilon^2)+R''_e(-m_\varepsilon^2)m_\varepsilon^2)} = \frac{C \sqrt{\varepsilon}}{\sqrt{N}} e^{-NR_e(-m_\varepsilon^2)}. \tag{2.121}
\]

**Step 3.** [Estimation of the term IV.] Using Jensen’s inequality, we have that \( \sum_{i=0}^{N-1} x_i^4 \geq N(Px)^4 \). Then, via the coarea formula,

\[
|IV| \leq \int_{\{x \in \mathbb{R}^N \mid Px < -R\}} e^{-\frac{1}{\varepsilon} \sum_{i=0}^{N-1} \frac{x_i}{2}y_i^2} \, e^{-\frac{1}{2}N\frac{1}{2}(Py)^4} \frac{1}{2} \, dy
\]

\[
= \sqrt{N} \int_{-\infty}^{-R} e^{-\frac{1}{2}N\frac{1}{2}m^4} \frac{1}{2} \, dm \int_{P^{-1}(m)} e^{-\frac{1}{2} \sum_{i=0}^{N-1} \frac{x_i}{2}y_i^2} \, d\mathcal{H}^{-1}. \tag{2.122}
\]

In Lemma C.3, we show that for all \( m \in \mathbb{R} \),

\[
\int_{P^{-1}(m)} e^{-\frac{1}{2} \sum_{i=0}^{N-1} \frac{x_i}{2}y_i^2} \, d\mathcal{H}^{-1} = e^{-N\frac{1}{2}(\frac{1}{2}m^2 + N\frac{1}{2}\log(2\varepsilon (J-1)^{-1}))} \sqrt{\frac{J-1}{\varepsilon^2}}. \tag{2.123}
\]

Therefore, by [10, 1.1], we have that for \( \varepsilon \) small enough,

\[
|IV| \leq \sqrt{N} \int_{-\infty}^{-R} e^{-\frac{1}{2}N\frac{1}{2}m^4} \frac{1}{2} \, dm \, \frac{\sqrt{J-1}}{\varepsilon^2} \leq \sqrt{N} \int_{-\infty}^{-R} e^{-\frac{1}{2}N\frac{1}{2}(\frac{R^2}{\varepsilon^2}-1)m^2} \, dm \, \sqrt{J-1} \varepsilon^2 \frac{1}{\varepsilon^2}
\]

\[
= C \int_{-\infty}^{-R} e^{-\frac{1}{2}m^2} \, dm \leq C \frac{\varepsilon}{\sqrt{N}} e^{-\frac{N\frac{1}{2}(\frac{R^2}{\varepsilon^2}-1)R^2}. \tag{2.124}
\]
Note that $\bar{H}_\varepsilon(-m^*_\varepsilon) \leq \frac{c}{\varepsilon}$ for some $c > 0$. Indeed, by Lemma C.1 (ii), we have that for some bounded function $\tau_\varepsilon$,
\begin{equation}
|\varphi_\varepsilon(-m^*_\varepsilon)| = \left| -\int_{-m^*_\varepsilon}^{0} \varphi'_\varepsilon(m) \, dm + \varphi_\varepsilon(0) \right| = \left| -\frac{1}{\varepsilon} \int_{-m^*_\varepsilon}^{0} \tau_\varepsilon(m) \, dm + \varphi_\varepsilon(0) \right| \leq \frac{1}{\varepsilon} \|\tau_\varepsilon\|_{L^\infty(K; dm)} m^*_\varepsilon + |\varphi_\varepsilon(0)|,
\end{equation}
and by (A.4) and (B.1),
\begin{equation}
\varphi_\varepsilon(0) = -\varphi^*_\varepsilon(0) = \log \int_{\mathbb{R}} e^{-\frac{1}{2}\psi_j(z)} \, dz \leq \frac{1}{2} \log (C\varepsilon).
\end{equation}
Combining (2.125) and (2.126) with the definition of $\bar{H}_\varepsilon$, shows that $\bar{H}_\varepsilon(-m^*_\varepsilon) \leq \frac{c}{\varepsilon}$ for some $c > 0$. Then, choosing $R$ large enough, (2.124) implies that
\begin{equation}
IV = O \left( \frac{1}{\sqrt{N}} \frac{e^{-N\bar{H}_\varepsilon(-m^*_\varepsilon)}}{\sqrt{\bar{H}''(m^*_\varepsilon)}} \sqrt{\varphi''_\varepsilon(-m^*_\varepsilon)} \right).
\end{equation}
This shows that the term $IV$ is of lower order than $I$, and concludes the proof. \hfill \Box

3 Rough estimates at high temperature

In this chapter, we consider the same system as in Chapter 2, but with two key differences. First, we do not consider the low-temperature regime here, that is, throughout this chapter, we suppose that $\varepsilon = 1$. The second difference is that, instead of $\psi(z) = z^4/4 - z^2/2$, we consider here a class of single-site potentials given by functions of the form $z \mapsto \Psi(z) - \frac{J}{2} z^2$, where $\Psi : \mathbb{R} \to \mathbb{R}$ satisfies Assumption 3.1 below.

Hence, the microscopic Hamiltonian $H^{N,1} : \mathbb{R}^N \to \mathbb{R}$ in this chapter is given by
\begin{equation}
H^{N,1}(x) = \sum_{i=0}^{N-1} \left( \Psi(x_i) - \frac{J}{2} x_i^2 \right) + \frac{J}{4N} \sum_{i,j=0}^{N-1} (x_i - x_j)^2 = \sum_{i=0}^{N-1} \Psi(x_i) - \frac{J}{2N} \sum_{i,j=0}^{N-1} x_i x_j,
\end{equation}
where $J > 0$ and we make the following assumptions on the single site potential $\Psi$.

Assumption 3.1 (1) There is a splitting $\Psi = \Psi_c + \Psi_b$ for some $\Psi_c, \Psi_b \in C^2(\mathbb{R})$, and there are constants $0 < c, C < \infty$ such that $\Psi''_c(z) \geq c$ and $|\Psi_b|_{C^2} \leq C$.

(2) $\Psi(z) = \Psi(-z)$ for all $z \in \mathbb{R}$.

(3) $z \mapsto \Psi'(z)$ is convex on $[0, \infty)$.

(4) If $\Psi_c$ is a quadratic function of the form $\Psi_c(x) = c_\Psi x^2 + c'_\Psi x + c''_\Psi$ for some $c_\Psi, c'_\Psi, c''_\Psi \in \mathbb{R}$, then we suppose that $c_\Psi > J$.

(5) $1/J < \int_{\mathbb{R}} z^2 e^{-\Psi(z)} \, dz / (\int e^{-\Psi(z)} \, dz)$.

(6) $\sigma \mapsto \int_{\mathbb{R}} (\Psi''(z))^2 e^{-\Psi(z) + \sigma z} \, dz / (\int e^{-\Psi(z) + \sigma z} \, dz)$ is locally bounded on $\mathbb{R}$. 

27
Remark 3.2 If $\Psi = \Psi_c$ is a quadratic function, then Assumption 3.1 is not fulfilled for any choice of $J$, and we do not expect to obtain Kramers’ law, since the macroscopic Hamiltonian $\bar{H}_1$ is not of double-well form, where $\bar{H}_1$ is defined as in (2.13) with $\psi$ being replaced by the function $z \mapsto \Psi(z) - \frac{J}{2}z^2$ and with $\varepsilon = 1$. Indeed, from (C.27), we see that $\bar{H}_1$ is a quadratic function and hence not of double-well form.

This chapter is organized similarly as Chapter 2. That is, in Section 3.1 we introduce the local Cramér theorem and show that the macroscopic Hamiltonian has a double-well structure. In Section 3.2 we formulate the main result of this chapter, which provides rough estimates on the average transition time between the metastable sets, where the metastable sets are defined analogously to Chapter 2. The only thing left to prove for this result is the lower bound on the capacity. This is done in Section 3.3.

3.1 Preliminaries

3.1.1 Local Cramér Theorem.

Replacing $\psi$ by the function $z \mapsto \Psi(z) - \frac{J}{2}z^2$ and setting $\varepsilon = 1$, we can proceed exactly as in (1.3) and Subsection 2.1.1 to define the Gibbs measure $\mu^{N,1}$ and to introduce a disintegration of $\mu^{N,1}(dx) = \mu_m^{N,1}(dx)\mu^{N,1}(dm)$. Analogously, we define the quantities $\varphi_{N,1}^1$, $\varphi_{1}^1$, $\varphi_1$ and $\mu_1^{\sigma}$ by (2.6), (2.8), (2.9) and (2.10), respectively, by replacing $\psi$ by the function $z \mapsto \Psi(z) - \frac{J}{2}z^2$ and setting $\varepsilon = 1$. Then, the local Cramér theorem in this chapter is given as follows.

**Proposition 3.3 (Local Cramér theorem)** Suppose Assumption 3.1. Then,

$$e^{-N\varphi_{N,1}(m)} = e^{-N\varphi_1(m)} \frac{\sqrt{\varphi_1''(m)}}{\sqrt{2\pi}} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right).$$

(3.2)

In particular,

$$d\mu^{N,1}(m) = \frac{1}{Z_{\mu^{N,1}}} e^{-NH_1(m)} \frac{\sqrt{\varphi_1''(m)}}{\sqrt{2\pi}} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right) dm.$$  

(3.3)

**Proof.** Using the same notation and proceeding as in the proof of Proposition C.2, we observe that it suffices to show that

$$\left|g_{N,m}(0) - \frac{1}{\sqrt{2\pi}}\right| = O\left(\frac{1}{\sqrt{N}}\right).$$

(3.4)

However, this was already shown in [25, Proposition 3.1 and Lemma 3.2].

3.1.1 Analysis of the energy landscape. In the following lemma we show that the macroscopic Hamiltonian $H_1$ has the form of a double-well function with at least quadratic growth at infinity.

**Lemma 3.4** Suppose Assumption 3.1. If $\Psi_c$ is a quadratic function, then let $c_{\Psi}$ denote the leading order coefficient. Otherwise, let $c_{\Psi} = \infty$. Then, we have that
(i) \( \liminf_{t \to \infty} \frac{\varphi_1(t)}{t^2} \geq c_\psi \), \( \liminf_{t \to \infty} \frac{B_1(t)}{t^2} \geq c_\psi - J/2 \)

(ii) there exists \( K_J > 0 \) and \( \delta > 0 \) such that \( \varphi_1'(t) \geq (J + \delta)t \) for all \( t \geq K_J \) and \( \varphi_1'(t) \leq (-J - \delta)t \) for all \( t \leq -K_J \), and

(iii) \( \dot{H}_1 \) has exactly three critical points located at \( -m_1^* \), \( 0 \) and \( m_1^* \) for some \( m_1^* > 0 \). Moreover, \( \dot{H}_1''(0) < 0 \), \( \dot{H}_1''(m_1^*) > 0 \) and \( \dot{H}_1''(-m_1^*) > 0 \). That is, \( \dot{H}_1 \) has a local maximum at \( 0 \), and the two global minima of \( \dot{H}_1 \) are located at \( \pm m_1^* \).

**Proof.** Since \( \varphi_1(t) = \varphi_1(-t) \) for all \( t \in \mathbb{R} \), it suffices to prove all claims only on \([0, \infty)\).

(i). As in Lemma 2.2, this statement follows from a simple argument given in [26, III.2.6].

(ii). From part (i) and Assumption 3.1 (4), we know that there exists \( K' > 0 \) and \( \delta' > 0 \) such that \( \varphi_1(t) \geq (J + \delta')t^2 \) for all \( t \geq K' \). Using that \( t \mapsto \varphi_1'(t) \) is increasing (since \( \varphi_1 \) is strictly convex) we obtain that for all \( t \geq K' \),

\[
(J + \delta')t^2 \leq \varphi_1(t) = \int_0^t \varphi_1'(r) dr + \varphi_1(0) \leq \varphi_1(t) + \varphi_1(0),
\]

which concludes the claim.

(iii). Before we show the claims, note that the function \( z \mapsto \varphi_1'(z) \) is convex on \([0, \infty)\).

Indeed, from [19, Theorem 1.2 c)], we know that Assumption 3.1 yields that \( z \mapsto (\varphi_1^*)'(z) \) is concave on \([0, \infty)\) (cf. [26, IV.0.4]). Hence, for \( w > z \), we have that \( (\varphi_1^*)''(\varphi_1'(w)) \leq (\varphi_1^*)''(\varphi_1'(z)) \), since, due to the convexity of \( \varphi_1 \), we have that \( \varphi_1'(w) \geq \varphi_1'(z) \). Therefore,

\[
\varphi_1''(w) = \frac{1}{(\varphi_1^*)''(\varphi_1'(w))} \geq \frac{1}{(\varphi_1^*)''(\varphi_1'(z))} = \varphi_1''(z),
\]

which shows that \( z \mapsto \varphi_1'(z) \) is convex.

To show that \( \dot{H}_1 \) admits a local maximum at \( 0 \), we observe that, since \( \varphi_1'(0) = 0 \), we have that \( \dot{H}_1'(0) = 0 \). Moreover, Assumption 3.1 implies that \( (\varphi_1^*)''(0) > 1/J \). Therefore, \( \varphi_1''(0) < J \) and \( \dot{H}_1''(0) < 0 \).

It remains to show that there exists a unique point \( m_1^* \in (0, \infty) \) such that \( \dot{H}_1'(m_1^*) = 0 \) and \( \dot{H}_1''(m_1^*) > 0 \). Using again that \( \varphi_1''(0) < J \), we infer that for \( z > 0 \) small enough,

\[
\varphi_1'(z) = \int_0^z \varphi_1''(r) dr < Jz.
\]

Moreover, by part (ii), we know that there exists \( m_1^* > z > 0 \) such that

\[
\varphi_1'(m_1^*) = Jm_1^* \quad \text{and} \quad \varphi_1'(z) < Jz \quad \text{for all} \quad z \in (0, m_1^*).
\]

However, the mean value theorem implies that there exists \( z' \in (0, m_1^*) \) such that \( \varphi_1''(z') > J \). Together with the fact that \( \varphi_1'' \) is non-decreasing, this implies that \( \varphi_1'(z) > J \) for all \( z \geq m_1^* \). This is in turn yields that

\[
\varphi_1'(z) > Jz \quad \text{for all} \quad z > m_1^* \quad \text{and} \quad \dot{H}_1''(m_1^*) > 0.
\]

Combining (3.8) and (3.9) shows that, at \( m_1^* \), there is the unique global minimum of \( \dot{H}_1 \) on \([0, \infty)\). \qed

29
3.2 Rough estimates on the average transition time

In this section we formulate the main result of this chapter. Most of its proof is omitted, since it is a straightforward adaptation from the proof of Theorem 2.7. However, for the proof of the lower bound on the capacity, we provide a new proof, which is given in Section 3.3.

**Theorem 3.5** Let $\pm m_1^*$ be the two global minimisers of the macroscopic Hamiltonian $\tilde{H}_1$. Let $\eta_1 > 0$ and $B_1^-, B_1^+ \subset \mathbb{R}^N$ be defined by (2.26) and (2.27) with $\varepsilon = 1$. Then, for some $a > 0$, which is independent of $N$,

$$
\mathbb{E}_{\nu_{B_1^-}}[T_{B_1^+}] \geq \frac{2\pi}{\sqrt{H''_1(-m_1^*)}} e^{\nu_1(B_1(0)-\tilde{H}_1(-m_1^*))} \left( 1 + O\left( \frac{\log(N)^3}{N} \right) \right), \quad \text{and} \quad (3.10)
$$

$$
\mathbb{E}_{\nu_{B_1^-}}[T_{B_1^+}] \leq (1 + a)\frac{2\pi}{\sqrt{H''_1(-m_1^*)}} e^{\nu_1(B_1(0)-\tilde{H}_1(-m_1^*))} \left( 1 + O\left( \frac{\log(N)^3}{N} \right) \right). \quad (3.11)
$$

**Proof.** As in the proof of Theorem 2.7, the starting point is the formula (2.22). Then, proceeding exactly as in the proofs of Proposition 2.8 and Proposition 2.14, we can show that

$$
\text{Cap}(B_1^-, B_1^+) \leq \frac{1}{2\pi} e^{-N\tilde{H}(0)} \sqrt{|H''_1(0)|} \left( 1 + O\left( \frac{\log(N)^3}{N} \right) \right), \quad \text{and} \quad (3.12)
$$

$$
\int_{(B_1^+)^c} f_{B_1^-}^*(y) e^{-\tilde{H}(y)} dy = e^{-N\tilde{H}_1(-m_1^*)} \sqrt{\varphi''_1(-m_1^*)} \left( 1 + O\left( \frac{\log(N)^3}{N} \right) \right), \quad (3.13)
$$

which yields (3.10). Finally, (3.11) follows from combining (3.13) with Proposition 3.6. This concludes the proof of this theorem. \qed

3.3 Rough lower bound on the capacity

In this section, we prove the rough lower bound on the capacity. The proof is inspired by the two-scale approach, which was initiated in [21]. To apply this approach, we use that, by [25, 1.6], for all $N \in \mathbb{N}$ and $m \in \mathbb{R}$, $\mu_{m,1}^{N,1}$ satisfies the Poincaré inequality with a constant $q > 0$, which is independent of $N$ and $m$. That is, for all $N \in \mathbb{N}$ and $m \in \mathbb{R}$ and for all $f \in H^1(\mu_{m,1}^{N,1})$,

$$
\text{Var}_{\mu_{m,1}^{N,1}}(f) := \int \left| f - \int f d\mu_{m,1}^{N,1} \right|^2 d\mu_{m,1}^{N,1} \leq \frac{1}{q} \int (i d - NP^t \nabla f)^2 d\mu_{m,1}^{N,1}, \quad (3.14)
$$

where $P^t m = (1/N)(m, \ldots, m) \in \mathbb{R}^N$ for $m \in \mathbb{R}$.

**Proposition 3.6** Consider the same setting as in Theorem 3.5. Let

$$
a = \frac{1}{\rho^2} \max_{m \in [-m_1^*, m_1^*]} \int \left| \Psi'' - \int \Psi'' d\mu_{1,\varphi_1(m)} \right|^2 d\mu_{1,\varphi_1(m)}. \quad (3.15)
$$

Then,

$$
\text{Cap}(B_1^-, B_1^+) \geq \frac{1}{1 + a} \frac{1}{2\pi} e^{-N\tilde{H}(0)} \sqrt{|H''_1(0)|} \sqrt{\varphi''_1(0)} \left( 1 + O\left( \frac{\log(N)^3}{N} \right) \right). \quad (3.16)
$$
Proof. Let \( f = f_{B_1^+, B_1^-}^*(x) \) and let

\[
\tilde{f}(m) = \int_{P^{-1}(m)} f \, d\mu_{m,1}^N.
\] (3.17)

As in [21, Section 2.1], we split the gradient \( \nabla f \) into its fluctuation part \((id - NP^t P)\nabla f\) and its macroscopic part \(NP^t P \nabla f\) and use that

\[
|(id - NP^t P)\nabla f|^2 + |NP^t P \nabla f|^2 = |\nabla f|^2.
\] (3.18)

Using the fact that \(|NP^t P x| = N|Px|\) for all \( x \in \mathbb{R}^N \), Jensen’s inequality and [21, Lemma 21], we obtain that

\[
\int |NP^t P \nabla f|^2 \, d\mu_{N,1}^N \geq N \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} \int_{P^{-1}(m)} |P \nabla f|^2 \, d\mu_{m,1}^N \, d\bar{\mu}_{N,1}(m)
\]
\[
\geq N \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} \int_{P^{-1}(m)} P \nabla f \, d\mu_{m,1}^N \bigg| ^2 \, d\bar{\mu}_{N,1}(m)
\]
\[
= N \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} \bigg| \frac{\tilde{f}(m)}{N} + PCov_{\mu_{m,1}^N}(f, \nabla H) \bigg|^2 \, d\bar{\mu}_{N,1}(m),
\] (3.19)

where, for two functions \( g, h \in L^1(\mu_{m,1}^N) \),

\[
Cov_{\mu_{m,1}^N}(g, h) = \int g \, (h - \int h \, d\mu_{m,1}^N) \, d\mu_{m,1}^N.
\] (3.20)

Then, using Young’s inequality, we have that, for some \( \tau \in (0,1) \),

\[
\int |NP^t P \nabla f|^2 \, d\mu_{N,1}^N \geq (1 - \tau) \frac{1}{N} \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} |\tilde{f}(m)|^2 \, d\bar{\mu}_{N,1}(m)
\]
\[
+ \left( 1 - \frac{1}{\tau} \right) N \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} \bigg| PCov_{\mu_{m,1}^N}(f, \nabla H) \bigg|^2 \, d\bar{\mu}_{N,1}(m).
\] (3.21)

Later in this proof we show that

\[
\frac{1}{N} \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} |\tilde{f}(m)|^2 \, d\bar{\mu}_{N,1}(m) \geq \frac{e^{-NH(0)}}{2\pi Z_{\mu_{N,1}}^N} \sqrt{\Theta_H(0)} \sqrt{\varepsilon_1''(0)} \left( 1 + O \left( \sqrt{\frac{\log(N)^3}{N}} \right) \right),
\] (3.22)

and

\[
\int_{m_1^*-\eta_1}^{m_1^*+\eta_1} \bigg| PCov_{\mu_{m,1}^N}(f, \nabla H) \bigg|^2 \, d\bar{\mu}_{N,1}(m) \leq \frac{\alpha N}{N} \int |(id - NP^t P)\nabla f|^2 \, d\mu_{m,1}^N \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right).
\] (3.23)

Combining (3.21), (3.22), (3.23) and (3.18) and optimizing in \( \tau \) yields (3.16). It only remains to show (3.22) and (3.23).

Proof of (3.22). Note that by Proposition 3.3,

\[
\frac{1}{N} \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} |\tilde{f}(m)|^2 \, d\bar{\mu}_{N,1}(m)
\]
\[
= \frac{1}{NZ_{\mu_{N,1}}^N} \int_{m_1^*-\eta_1}^{m_1^*+\eta_1} |\tilde{f}|^2 \frac{\sqrt{\varepsilon_1''(m)}}{2\pi} e^{-NR_1(m)} \, dm \left( 1 + O \left( \sqrt{\frac{\log(N)^3}{N}} \right) \right).
\] (3.24)
Then, by our knowledge on one-dimensional capacities (see for instance [11, Section 7.2.5]),
\[
\int_{\frac{m_1 - n_1}{m_1 + n_1}}^{\frac{m_1 - n_1}{m_1 + n_1}} |f'|^2 \frac{\sqrt{\phi_1'(m)}}{\sqrt{2\pi}} e^{-NH_1(m)} dm \\
\geq \inf_{h: h(-m_1 + n_1) = 1, h(m_1 - n_1) = 0} \int_{\frac{m_1 - n_1}{m_1 + n_1}}^{\frac{m_1 - n_1}{m_1 + n_1}} |h'|^2 \frac{\sqrt{\phi_1'(m)}}{\sqrt{2\pi}} e^{-NH_1(m)} dm \\
= \frac{1}{\sqrt{2\pi}} \left( \int_{\frac{m_1 - n_1}{m_1 + n_1}}^{\frac{m_1 - n_1}{m_1 + n_1}} \sqrt{\phi_1'(r)} e^{N\hat{H}_1(r)} dr \right)^{-1}.
\]
(3.25)
Recalling that \( \max_{m\in[-m_1+n_1,m_1-n_1]} H_1(m) = \hat{H}_1(0) \) and \( \sqrt{N}Z_{\mu,N,1} = Z_{\mu,N,1} \) by the coarea formula, we conclude (3.22) from standard Laplace asymptotics.

**Proof of (3.23).** Since \( \mu_{m,N}^{1} \) is supported on \( P^{-1}(m) \), we have that
\[
P\text{Cov}_{\mu_{m,N}^{1}}(f, \nabla H) = \frac{1}{N} \text{Cov}_{\mu_{m,N}^{1}} \left( f, \sum_{i=0}^{N-1} \Psi'(x_i) \right),
\]
Then, using Hölder’s inequality and (3.14),
\[
\left| P\text{Cov}_{\mu_{m,N}^{1}}(f, \nabla H) \right|^2 \leq \frac{1}{N^2} \text{Var}_{\mu_{m,N}^{1}}(f) \text{ Var}_{\mu_{m,N}^{1}} \left( \sum_{i=0}^{N-1} \Psi'(x_i) \right) \\
\leq \frac{1}{\vartheta^2 N^2} \int \left| (id - NP^t P) \nabla f \right|^2 d\mu_{m,N}^{1} \int \left| (id - NP^t P) \nabla \sum_{i=0}^{N-1} \Psi'(x_i) \right|^2 d\mu_{m,N}^{1}.
\]
(3.26)
It remains to show that the second term on the right-hand side of (3.26) is bounded linearly in \( N \). To do this, we use that by Lemma 3.7 and Corollary 3.8,
\[
\int \left| (id - NP^t P) \nabla \sum_{i=0}^{N-1} \Psi'(x_i) \right|^2 d\mu_{m,N}^{1} = \int \sum_{i=0}^{N-1} \left| \Psi''(x_i) - \frac{1}{N} \sum_{j=1}^{N} \Psi''(x_j) \right|^2 d\mu_{m,N}^{1} \\
= (N - 1) \int \left| \Psi'' - \int \Psi'' d\mu_{m,N}^{1,\varphi'_i(m)} \right|^2 d\mu_{m,N}^{1,\varphi'_i(m)} \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right) \\
\leq N \max_{m\in[-m_1,n_1]} \int \left| \Psi'' - \int \Psi'' d\mu_{m,N}^{1,\varphi'_i(m)} \right|^2 d\mu_{m,N}^{1,\varphi'_i(m)} \left( 1 + O \left( \frac{1}{\sqrt{N}} \right) \right) \\
\]
(3.27)
This concludes the proof of (3.23).

**3.3.1 Equivalence of observables**

In (3.27) we have used that we can pass from expectations with respect to \( \mu_{m,N}^{1} \), to expectations with respect to \( \otimes_{i=1}^{N} \mu^{1,\varphi'_i(m)} \). Such a statement is known in the literature as the **equivalence of observables** (see [23]). In Lemma 3.7, we show this result rigorously for functions that only depend on one coordinate. However, a straightforward modification of the proof yields the claim also for **local functions**, i.e. for functions that depend only on finitely many coordinates, see Corollary 3.8. The proof combines the ideas from [23] and [25].
Lemma 3.7 Let $b : \mathbb{R} \to [0, \infty)$ be such that
\begin{equation}
\sup_{m \in [-m^*_1, m^*_1]} \int_{\mathbb{R}} b(z) \, d\mu^{1,\varphi'_1(m)}(z) < \infty. \tag{3.28}
\end{equation}
Then, there exists $C_b \in (0, \infty)$, independent of $m \in [-m^*_1, m^*_1]$ and $N \in \mathbb{N}$, such that, for all $m \in [-m^*_1, m^*_1]$ and $N$ large enough,
\begin{equation}
\left| \int_{p^{-1}(m)} b(x_1) \, d\mu^{N,1}_m - \int_{\mathbb{R}} b(z) \, d\mu^{1,\varphi'_1(m)}(z) \right| \leq C_b \frac{1}{\sqrt{N}}. \tag{3.29}
\end{equation}

Proof. Fix $m \in [-m^*_1, m^*_1]$. In this proof, let $C$ denote a varying positive constant, which does not depend on $N$ and $m$, but may depend in $b$.

Let $\mu^{1,\varphi'_1(m),N} = \otimes_{i=1}^N \mu^{1,\varphi'_1(m)}$. Proceeding as in [23], for $\sigma \in \mathbb{R}$, we define
\begin{equation}
\mu^{\sigma,1,\varphi'_1(m),N}(dx) = \frac{1}{Z} \exp \left( \varphi'_1(m) \sum_{i=0}^{N-1} x_i + \sigma b(x_1) - \sum_{i=0}^{N-1} \Psi(x_i) \right) \, dx, \tag{3.30}
\end{equation}
where $Z$ denotes the normalization constant. Note that $\mu^{0,1,\varphi'_1(m),N} = \mu^{1,\varphi'_1(m),N}$. Let $(Y_i)_{i=1,\ldots,N}$ be a random vector distributed according to $\mu^{\sigma,1,\varphi'_1(m),N}$. Let
\begin{equation}
S_{N,m} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} (Y_i - m), \tag{3.31}
\end{equation}
and let $\tilde{g}_0$ be the Lebesgue density of the distribution of $S_{N,m}$. Using the same arguments as in [23, 4.1 and 4.2], we observe that
\begin{equation}
\left| \int_{\mathbb{R}^N} b(x_1) \, d\mu^{1,\varphi'_1(m),N} - \int_{p^{-1}(m)} b(x_1) \, d\mu^{N,1}_m \right| = - \frac{d}{d\sigma} \frac{\tilde{g}_0(\sigma)}{\tilde{g}_0(0)}, \tag{3.32}
\end{equation}
Let $s(m) = \varphi''_1(m)^{-1}$. In the proof of Proposition 3.3, we have seen that
\begin{equation}
\tilde{g}_0(0) = \exp (N \varphi_1(m) - N \varphi_{1,1}(m)) = \frac{1}{s(m)\sqrt{2\pi}} + O \left( \frac{1}{N} \right) \geq C, \tag{3.33}
\end{equation}
since $s$ is bounded from above and from below uniformly in $m \in [-m^*_1, m^*_1]$. Hence,
\begin{equation}
\left| \int_{\mathbb{R}^N} b(x_1) \, d\mu^{1,\varphi'_1(m),N} - \int_{p^{-1}(m)} b(x_1) \, d\mu^{N,1}_m \right| \leq C \left| \frac{d}{d\sigma} \frac{\tilde{g}_0(\sigma)}{\tilde{g}_0(0)} \right|. \tag{3.34}
\end{equation}

For any function $f : \mathbb{R} \to \mathbb{R}$ and for all $x \in \mathbb{R}$, we abbreviate
\begin{equation}
\langle f \rangle := \int_{\mathbb{R}} f(z) \, d\mu^{1,\varphi'_1(m)}(z) \quad \text{and} \quad \hat{x} = \hat{x}(x) = \frac{x - m}{s}. \tag{3.35}
\end{equation}

Let $(X_i)_{i=1,\ldots,N}$ be a random vector distributed according to $\mu^{1,\varphi'_1(m),N}$, and let $X$ be a random variable distributed according to $\mu^{1,\varphi'_1(m)}$. By [23, 4.4], we have that
\begin{equation}
2\pi \frac{d}{d\sigma} \frac{\tilde{g}_0(\sigma)}{\tilde{g}_0(0)} = \int_{\mathbb{R}} \mathbb{E}_{\mu^{1,\varphi'_1(m),N}} \left[ (b(X_1) - \langle b \rangle)e^{i \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} (X_1 - m)\xi} \right] d\xi
= \int_{\mathbb{R}} \mathbb{E}_{\mu^{1,\varphi'_1(m)}} \left[ (b(X) - \langle b \rangle)e^{i \frac{1}{\sqrt{N}} (X-m)\xi} \right] \mathbb{E}_{\mu^{1,\varphi'_1(m)}} \left[ e^{i \frac{1}{\sqrt{N}} (X-m)\xi} \right]^{N-1} d\xi
= s(m)^{-1} \int_{\mathbb{R}} \left( (b - \langle b \rangle)e^{i \frac{1}{\sqrt{N}} \hat{x}(m)\xi} \right) \left( e^{i \frac{1}{\sqrt{N}} \hat{x}(m)\xi} \right)^{N-1} d\xi. \tag{3.36}
\end{equation}
Moreover, by Taylor’s formula, for some $\theta$

$$\left| \int_{\mathbb{R}^N} b(x_1) \, d\mu^1_{x_1(m),N} - \int_{P^{-1}(m)} b(x_1) \, d\mu_{m,1} \right| \leq C \left| \int_{\mathbb{R}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi \right|^{N-1} d\xi. \quad (3.37)$$

It remains to show that for $N$ large enough,

$$\left| \int_{\mathbb{R}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi \right|^{N-1} d\xi = O \left( \frac{1}{\sqrt{N}} \right). \quad (3.38)$$

In order to show (3.38), we proceed as in the proof of [25, 3.1] (see also Proposition C.2). As it was shown there, there exists $\hat{\delta} > 0$, $C > 0$ and a complex-valued function $h$ such that (see [25, (46)]) for all $|\xi| \leq \hat{\delta}$,

$$\left| e^{i\xi^2} \right| = e^{-h(\xi)} \quad \text{and} \quad \left| h(\hat{\xi}) - \frac{1}{2} \hat{\xi}^2 \right| \leq C|\hat{\xi}|^3. \quad (3.39)$$

We split the integral on the left-hand side in (3.38) according to some $\delta < \hat{\delta}$ (which we choose later) such that

$$\int_{\{|\frac{\xi}{\sqrt{N}}| \leq \delta\}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi \quad \text{and} \quad \int_{\{|\frac{\xi}{\sqrt{N}}| > \delta\}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi$$

$$=: I + II. \quad (3.40)$$

To estimate the term $I$, we argue similarly as in [25, (68)] to observe that

$$\left| (N - 1)h \left( \frac{\xi}{\sqrt{N}} \right) - \frac{1}{2} \hat{\xi}^2 \right| \leq C \left( \frac{|\hat{\xi}|^3}{\sqrt{N}} + \frac{|\hat{\xi}|^2}{2N} \right). \quad (3.41)$$

As in [25, (69) and p.32], this inequality yields that (for $\delta$ small enough and $N$ large enough),

$$\left| e^{(N-1)h \left( \frac{\xi}{\sqrt{N}} \right)} - e^{-\frac{1}{2} \hat{\xi}^2} \right| \leq e^{-\frac{1}{2} \hat{\xi}^2} C \left( \frac{|\hat{\xi}|^3}{\sqrt{N}} + \frac{|\hat{\xi}|^2}{N} \right). \quad (3.42)$$

This implies that

$$\left| I - \int_{\{|\frac{\xi}{\sqrt{N}}| \leq \delta\}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi \right| \leq C \int_{\{|\frac{\xi}{\sqrt{N}}| \leq \delta\}} \left( \frac{|\hat{\xi}|^3}{\sqrt{N}} + \frac{|\hat{\xi}|^2}{N} \right) e^{-\frac{1}{2} \hat{\xi}^2} d\xi \leq C \frac{1}{\sqrt{N}}. \quad (3.43)$$

since

$$\left| \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \right| \leq \sup_{m \in [-m^1, m^1]} \langle |b - \langle b \rangle| \rangle \leq \infty. \quad (3.44)$$

Moreover, by Taylor’s formula, for some $\theta \in [0, 1]$,

$$\int_{\{|\frac{\xi}{\sqrt{N}}| \leq \delta\}} \left( b - \langle b \rangle \right) e^{i\frac{x_1}{\sqrt{N}}} \, d\xi$$

$$\leq 0 + \frac{1}{\sqrt{N}} \int_{\{|\frac{\xi}{\sqrt{N}}| \leq \delta\}} \hat{x}(b - \langle b \rangle) e^{i\theta \frac{x_1}{\sqrt{N}}} \, d\xi \leq C \frac{1}{\sqrt{N}}, \quad (3.45)$$

$$34$$
which shows that $|I| = O \left( 1/\sqrt{N} \right)$.

Finally, to estimate the term $II$, we first use (3.44), and then use the same arguments as in [25, p. 32] for the integral

$$\int_{\{\frac{\xi}{\sqrt{N}} \geq \delta\}} \left( e^{\frac{i}{\sqrt{N}}} \right)^{N-1} d\xi. \quad (3.46)$$

This concludes the proof of (3.38).

A straightforward modification of the proof yields to the following extension of Lemma 3.7 to local functions.

**Corollary 3.8** Let $\ell \in \mathbb{N}$, and let $b : \mathbb{R}^\ell \to [0, \infty)$ be such that

$$\sup_{m \in [-m_1^*, m_1^*]} \int_{\mathbb{R}^\ell} b(z_1, \ldots, z_\ell) d\mu_{1,\ell}^{1,\varphi_1(m),\ell}(z_1, \ldots, z_\ell) < \infty, \quad (3.47)$$

where $\mu_{1,\ell}^{1,\varphi_1(m),\ell} = \otimes_{i=1}^{\ell} \mu_{1,\ell}^{1,\varphi_1(m)}$. Then there exists $C_b \in (0, \infty)$, independent of $m \in [-m_1^*, m_1^*]$ and $N \in \mathbb{N}$, such that, for all $m \in [-m_1^*, m_1^*]$ and $N$ large enough,

$$\left| \int_{\mathbb{P}_N} b(x_1, \ldots, x_N) d\mu_{N,\ell}^{1,\varphi_1(m),\ell}(x_1, \ldots, x_N) - \int_{\mathbb{R}^\ell} b(z_1, \ldots, z_\ell) d\mu_{1,\ell}^{1,\varphi_1(m),\ell}(z_1, \ldots, z_\ell) \right| \leq C_b \frac{1}{\sqrt{N}}. \quad (3.48)$$

**Appendix**

This appendix is organized as follows. In Section A we collect several properties of Cramér transforms and the cumulant generating functions. In Section B, we derive asymptotic expressions for integrals of the form

$$\int_{\mathbb{R}} f(z) e^{\frac{-1}{\varepsilon} U(z)} dz \quad (4.9)$$

for certain functions $f, U \in C^\infty(\mathbb{R})$ and for $\varepsilon \ll 0$. The methods are based on standard Laplace asymptotics. Finally, in Section C we state and prove the local Cramér theorem.

**A Properties of the Cramér transform**

**Lemma A.1** Let $W \in C^\infty(\mathbb{R})$ be such that $\liminf_{|z| \to \infty} W''(z) > 0$. Let

$$\chi^*(\sigma) = \log \int_{\mathbb{R}} e^{\sigma z - W(z)} dz, \quad \text{for } \sigma \in \mathbb{R}, \quad (A.1)$$

and let $\chi$ denote its Legendre transform, i.e.

$$\chi(m) = \sup_{\sigma \in \mathbb{R}} (\sigma m - \chi^*(\sigma)). \quad (A.2)$$

For all $\sigma \in \mathbb{R}$, define $\mu^\sigma \in \mathbb{P}(\mathbb{R})$ by

$$d\mu^\sigma(z) = e^{-\chi^*(\sigma)} e^{\sigma z - W(z)} dz = \frac{e^{\sigma z - W(z)}}{\int_{\mathbb{R}} e^{\sigma z - W(z)} dz} dz. \quad (A.3)$$

Then, the following statements hold true.
(i) $\chi^*$ and $\chi$ are strictly convex and smooth. If $W$ is even, then $\chi^*$ and $\chi$ are also even.

(ii) For $m \in \mathbb{R}$, we have that

$$\chi(m) = \chi'(m)m - \chi^*(\chi'(m)) \quad \text{and} \quad (\chi^*)'(\chi'(m)) = m. \quad \text{(A.4)}$$

In particular,

$$\chi''(m) = \frac{1}{(\chi^*)''(\chi'(m))} \quad \text{and} \quad \chi'''(m) = \frac{-(\chi^*)'''(\chi'(m))}{(\chi^*)''(\chi'(m))^3}. \quad \text{(A.5)}$$

(iii) For all $\sigma \in \mathbb{R}$,

$$\langle \chi^* \rangle'(\sigma) = \int_{\mathbb{R}} z e^{\sigma z - W(z)} \, dz = \int_{\mathbb{R}} z \, d\mu^\sigma(z),$$

$$\langle \chi^* \rangle''(\sigma) = \int_{\mathbb{R}} (z - \langle \chi^* \rangle'(\sigma))^2 \, d\mu^\sigma(z), \quad \text{(A.6)}$$

$$\langle \chi^* \rangle'''(\sigma) = \int_{\mathbb{R}} (z - \langle \chi^* \rangle'(\sigma))^3 \, d\mu^\sigma(z),$$

$$\langle \chi^* \rangle^{(4)}(\sigma) + 3 \langle \chi^* \rangle''(\sigma)^2 = \int_{\mathbb{R}} (z - \langle \chi^* \rangle'(\sigma))^4 \, d\mu^\sigma(z).$$

Proof. These are standard results that follow from some elementary computations. We refer to [26, III.2.5] and [21, Lemma 41] for more details.

\[ \square \]

B Some asymptotic integrals

The main result in this section is the following lemma, which is based on Laplace asymptotics. In the proof we use the same strategy as in [22, A.3].

Lemma B.1 Let $K \subset \mathbb{R}^2$ be a compact set. Let $U \in C^{0,\infty}(K \times \mathbb{R})$, and for $r \in K$, let $U_r(z) = U(r, z)$. Suppose that there exists $\alpha > 0$ and $R > 0$ such that, for all $r \in K$, $U_r$ admits a unique global minimum at some point $z_r \in \mathbb{R}$ with $U''_r(z_r) > R^{-1}$ and such that $U_r(z) \geq \alpha z^2$ for all $z \in [-R, R]^c$. Furthermore, we assume that the map $r \mapsto z_r$ is bounded on $K$. Then, for each $k \in \mathbb{N}_0$ and for each $r \in K$,

$$\int_{\mathbb{R}} (z - z_r)^{2k} e^{-\frac{1}{2}U_r(z)} \, dz = e^{-\frac{1}{2}U_r(z_r)} \frac{2\pi (2k - 1)!! \varepsilon^{k + \frac{5}{2}}}{U''_r(z_r)^{k + \frac{3}{2}}} \left( 1 + O_K \left( \sqrt{\varepsilon \log(\varepsilon^{-1})} \right) \right), \quad \text{(B.1)}$$

where for $n \in \mathbb{N}$, $n!!$ denotes the double factorial, and we make the convention that $(-1)!! := 1$.

Moreover,

$$\int_{\mathbb{R}} (z - z_r)^{2k+1} e^{-\frac{1}{2}U_r(z)} \, dz = -e^{-\frac{1}{2}U_r(z_r)} \frac{2\pi (2k + 1)!! U''_r(z_r) \varepsilon^{k + \frac{5}{2}}}{6U''_r(z_r)^{k + \frac{3}{2}}} \left( 1 + O_K \left( \sqrt{\varepsilon \log(\varepsilon^{-1})} \right) \right). \quad \text{(B.2)}$$
Proof. Fix \( r \in \mathcal{K} \) In this proof, let \( C \) denote a varying positive constant, which is independent of \( \varepsilon \) and \( r \).

**Step 1.** [Proof of (B.1).]

Let \( \rho = \sqrt{2(k + 1)\varepsilon \log(\varepsilon^{-1})/\bar{U}_r''(z_\varepsilon)} \) and \( \bar{U}_r(z) = U_r(z + z_\varepsilon) \). Let \( \bar{R} \geq R + \sup_{r \in \mathcal{K}} (|z_r| + \sqrt{|\bar{U}_r(z_r)|/\bar{a}^n}) \) be such that, for some \( \iota > 0 \), \( y^{2k} \leq e^{\iota U_r(y)} \) for all \( y \in [-\bar{R}, \bar{R}]^c \). Then,

\[
\int_{\mathbb{R}} (z - z_\varepsilon) e^{-\frac{2k}{y} U_r(z)} dz = \int_{\mathbb{R}} y^{2k} e^{-\frac{1}{y} U_r(y)} dy
\]

\[
= \int_{-\rho}^{\rho} y^{2k} e^{-\frac{1}{y} U_r(y)} dy + \int_{B_\bar{R}(0)^c} y^{2k} e^{-\frac{1}{y} U_r(y)} dy + \int_{B_\bar{R}(0) \setminus B_{\rho}(0)} y^{2k} e^{-\frac{1}{y} U_r(y)} dy
\]

\[
=: I + II + III.
\]

In the following we show that \( I \) provides the main contribution and that \( II \) and \( III \) are negligible.

**Step 1.1.** [Estimation of the term \( I \)]

Note that by Taylor’s formula, for some \( \theta \in [0, 1] \),

\[
\bar{U}_r(y) = U_r(z_\varepsilon) + \frac{1}{2} y^2 U_r''(z_\varepsilon) + \frac{1}{6} y^3 U_r'''(\theta y).
\]

By using that \( \bar{U}_r''' \) is locally bounded (uniformly in \( r \in \mathcal{K} \)), we see that there exists some \( c > 0 \) such that \( |U_r'''(\theta y)| \leq c \) for all \( y \in [-\rho, \rho] \). Therefore,

\[
e^{-\rho^3} \leq \int_{-\rho}^{\rho} y^{2k} e^{-\frac{1}{y} U_r(y)} dy \leq e^{\rho^3}.
\]

This, by using the definition of \( \rho \) and by some standard Gaussian computations applied to the denominator in (B.5), we infer that

\[
I = \sqrt{\frac{2\pi \varepsilon}{\bar{U}_r''(z_\varepsilon)}} e^{-\frac{1}{y} U_r(z_\varepsilon)} \left( \varepsilon^k \frac{(2k - 1)!!}{U_r''(z_\varepsilon)^k} + O_K \left( \varepsilon^{k+\frac{1}{2}} \sqrt{\log(\varepsilon^{-1})^3} \right) \right).
\]

**Step 1.2.** [Estimation of the term \( II \)]

We know that \( \bar{U}_r(y) \geq \alpha y^2 \) and \( y^{2k} \leq e^{\iota U_r(y)} \) for all \( y \in [-\bar{R}, \bar{R}]^c \). Hence, by [10, 1.1],

\[
II \leq 2 \int_{\bar{R}}^{\infty} e^{-\left(\frac{2}{\varepsilon} - 1\right) y^2} dy \leq C e^{-\frac{2}{\varepsilon} \bar{R}^2}.
\]

Since \( \alpha \bar{R}^2 > |U_r(z_\varepsilon)| \), this shows that

\[
II = e^{-\frac{1}{y} U_r(z_\varepsilon)} O_K \left( e^{k+\frac{1}{2}} \sqrt{\log(\varepsilon^{-1})^3} \right).
\]

**Step 1.3.** [Estimation of the term \( III \)]

Since \( \bar{U}_r \) has its unique minimum in \( 0 \), we have that for \( \varepsilon \) small enough, \( \inf_{y \in B_\bar{R}(0) \setminus B_{\rho}(0)} \bar{U}_r(y) = \bar{U}_r(\rho) \wedge \bar{U}_r(-\rho) \). Without restriction, we suppose that \( \bar{U}_r(\rho) \leq \bar{U}_r(-\rho) \). Then, by using (B.4) and the arguments from Step 1.1,

\[
III \leq 2 \bar{R} e^{\frac{1}{2} U_r(z_\varepsilon + \rho)} \leq C e^{-\frac{1}{2} U_r(z_\varepsilon)} e^{-\frac{1}{2} U_r''(z_\varepsilon) \frac{1}{2} \rho^2}.
\]

37
Using the definition of $\rho$, we have shown that
\[ III = e^{\frac{1}{2}U_t(z_t)} O_K \left( \varepsilon^{k+1} \sqrt{\log(\varepsilon^{-1})^3} \right). \]  

**Step 2.** [Proof of (B.2).]

(B.2) follows by proceeding exactly as in Step 1 (with $\bar{\rho} = \sqrt{2(2k+1)\varepsilon \log(\varepsilon^{-1})}/\sqrt{U_t''(z_t)}$ replacing $\rho$) but with the only difference that here we estimate the leading order term $I$ in the following way. The idea is based on Step 2.3 of the proof of [22, A.3]. First, by adding one more term in the Taylor expansion in (B.4), we have that for some $\theta \in [0,1]$,
\[ \bar{U}_t(y) = U_t^0 + \frac{1}{2} y^2 U_t^2 + \frac{1}{6} y^3 U_t^3 + \frac{1}{24} y^4 \bar{U}_t^4(\theta y), \]  

where for $i = 0, 1, 2, 3$, we abbreviate $U_t^i := U_t^{(i)}(z_t)$. Then,
\[ e^{\frac{1}{2}U_t(z_t)} I = e^{\frac{1}{2}U_t^0} \int_{-\bar{\rho}}^{\bar{\rho}} y^{2k+1} e^{-\frac{1}{2} U_t(y)} dy = \int_{-\bar{\rho}}^{\bar{\rho}} y^{2k+1} e^{-\frac{1}{2} (\frac{y^2}{2} U_t^2 + \frac{y^4}{24} U_t^4(\theta y))} dy 
- \frac{1}{6\varepsilon} U_t^3 \int_{-\bar{\rho}}^{\bar{\rho}} y^{2k+4} e^{-\frac{1}{2} y^2 U_t^2} dy - \frac{1}{24\varepsilon} \int_{-\bar{\rho}}^{\bar{\rho}} y^{2k+5} \bar{U}_t^4(\theta y) e^{-\frac{1}{2} \frac{y^2}{2} U_t^2} dy 
+ \int_{-\bar{\rho}}^{\bar{\rho}} y^{2k+1} e^{-\frac{1}{2} \frac{y^2}{2} U_t^2} \left( e^{-\frac{1}{2} \frac{y^2}{2} U_t^2 + \frac{y^4}{24} U_t^4(\theta y))} - 1 + \frac{y^3}{6\varepsilon} U_t^3 + \frac{y^4}{24\varepsilon} \bar{U}_t^4(\theta y) \right) dy 
=: I_1 + I_2 + I_3. \]

We now show that the term $I_1$ provides the dominant contribution and that $I_2$ and $I_3$ are of lower order than $I_1$. Concerning $I_1$, simple Gaussian computations as in Step 1.1 yield that
\[ I_1 = -\frac{1}{6} U_t^3 \sqrt{\frac{2\pi\varepsilon}{U_t''(z_t)}} \left( \varepsilon^{k+1} \frac{(2k+3)!}{U_t''(z_t)^{k+2}} + O_K \left( \varepsilon^{k+1} \frac{1}{\sqrt{\log(\varepsilon^{-1})}} \right) \right). \]  

For $I_2$ we use that $\bar{U}_t^4$ is locally bounded to obtain that
\[ |I_2| \leq \frac{C}{\varepsilon} \int_{-\bar{\rho}}^{\bar{\rho}} |y|^{2k+5} e^{-\frac{1}{2} \frac{y^2}{2} U_t^2} dy \leq C \varepsilon^{k+2}. \]  

Finally, to estimate the term $I_3$, note that $y^4 \leq y^3$ for $y \in [-\bar{\rho}, \bar{\rho}]$, $U_t'''$ and $U_t^4$ are locally bounded, and that $\bar{\rho}^3/\varepsilon \leq C \sqrt{\log(\varepsilon^{-1})}$. Then, by using the inequality $|e^{-x} - 1 + x| \leq |x|^2 e^{|x|}$,
\[ |I_3| \leq \frac{C}{\varepsilon^2} \int_{-\bar{\rho}}^{\bar{\rho}} |y|^{2k+1} e^{-\frac{1}{2} \frac{y^2}{2} U_t^2} \left| y^3 U_t^3 + \frac{y^4}{24\varepsilon} \bar{U}_t^4(\theta y) \right|^2 dy \leq \frac{C}{\varepsilon^2} \int_{-\bar{\rho}}^{\bar{\rho}} |y|^{2k+1} e^{-\frac{1}{2} \frac{y^2}{2} U_t^2} |y|^6 dy \leq C \varepsilon^{k+2}. \]

This concludes the proof of (B.2). \[ \square \]

**Corollary B.2** Consider the same setting as in Lemma B.1. Then,
\[ \frac{\int_{\mathbb{R}} (z - z_t)^{2k} e^{-\frac{1}{2}U_t(z)} dz}{\int_{\mathbb{R}} e^{-\frac{1}{2}U_t(z)} dz} = \varepsilon^k \frac{(2k-1)!}{U_t''(z_t)} + O_K \left( \varepsilon^{k+\frac{1}{2}} \sqrt{\log(\varepsilon^{-1})^3} \right), \quad \text{and} \]  
\[ \frac{\int_{\mathbb{R}} (z - z_t)^{2k+1} e^{-\frac{1}{2}U_t(z)} dz}{\int_{\mathbb{R}} e^{-\frac{1}{2}U_t(z)} dz} = -\frac{(2k+3)!}{6 U_t''(z_t)^{k+2}} + O_K \left( \varepsilon^{k+\frac{1}{2}} \sqrt{\log(\varepsilon^{-1})^3} \right). \]  

38
and suppose that there exists \( c > 0 \) such that \( \frac{1}{t} \leq c \) for all \( t \in I \). Notice that the inverse \((V_t')^{-1}\) of \( V_t'\) exists for all \( t \). Let \( \chi_{t,t}^\varepsilon, \chi_{t,t}^\varepsilon \) and \( \mu^\varepsilon, \sigma^\varepsilon \) be given by (A.1), (A.2) and (A.3), respectively, with \( W \) replaced by \( \varepsilon V_t \).
(i) Let \( \tilde{K} \subset \mathbb{R} \) be compact. Then, for all \((\lambda, t) \in \tilde{K} \times I\), \( (\chi_{\varepsilon,t})' \left( \frac{1}{\varepsilon} \right) = (V_t')^{-1}(\lambda) + \Omega_{\tilde{K} \times I}(\varepsilon) \).

(ii) For all compact intervals \( K \subset \mathbb{R} \) and all \( t \in I \) there exists a function \( \tau_{\varepsilon,t} : K \to \mathbb{R} \) and \( \varepsilon > 0 \) such that \( \sup_{0 < \varepsilon < \varepsilon_0} \sup_{(m,t) \in K \times I} |\tau_{\varepsilon,t}(m)| < \infty \) and \( \chi_{\varepsilon,t}(m) = \frac{1}{\varepsilon} \tau_{\varepsilon,t}(m) \) for all \((m,t) \in K \times I \) and for all \( \varepsilon < \varepsilon_0 \).

(iii) For \( m \in \mathbb{R} \) and \( t \in I \), let
\[
 s_t(m) := \left( \chi_{\varepsilon,t}^* \right)'(\chi_{\varepsilon,t}(m))^{\frac{1}{2}}. \tag{C.1}
\]

Then, for each compact interval \( K \subset \mathbb{R} \), there exist \( C_I > 0 \) and \( \varepsilon_0 > 0 \) such that for all \((m,t) \in K \times I \) and for all \( \varepsilon < \varepsilon_0 \),
\[
 s_t(m)^2 = \Omega_{K \times I}(\varepsilon) \quad \text{and} \quad \sum_{k=1}^{4} \int_{\mathbb{R}} \left| \frac{z - m}{s_t(m)} \right|^k d\mu_{\varepsilon}^{\varepsilon,t,\chi_{\varepsilon,t}(m)}(z) \leq C_I. \tag{C.2}
\]

**Proof.** (i). Note that for all \( \lambda \in \tilde{K} \), the function \( U((\lambda, t), z) = V_t(z) - \lambda z \) satisfies the same conditions as the function \( U \) from Corollary B.2. In particular, \( U(\lambda,t) \) admits a unique global minimum at \((V_t')^{-1}(\lambda)\). Thus, part (i) follows immediately from Lemma A.1 and (B.17).

(ii). Let \( K = [a, b] \) for some \( a, b \in \mathbb{R} \) with \( a < b \). Set \( F(m) = (\chi_{\varepsilon,t}^*)'(V_t'(m)/\varepsilon) \). From part (i), we know that for \( \varepsilon \) small enough,
\[
 F(a - 1) = a - 1 + \Omega_{[a-1,b+1] \times I}(\varepsilon) < a, \quad \text{and} \quad F(b + 1) = b + 1 + \Omega_{[a-1,b+1] \times I}(\varepsilon) > b. \tag{C.3}
\]

Therefore, by the continuity of \( F \) and the mean value theorem, \( F([a - 1, b + 1]) \supset K \). We also know that \( F : [a - 1, b + 1] \to F([a - 1, b + 1]) \) is bijective, since \( F \) is strictly increasing. Setting now \( \tau_{\varepsilon,t}(m) = V_t'(F^{-1}(m)) \) for \( m \in K \) yields that
\[
 (\chi_{\varepsilon,t}^*)' \left( \frac{1}{\varepsilon} \tau_{\varepsilon,t}(m) \right) = m \quad \text{for all} \ (m,t) \in K \times I. \tag{C.4}
\]

Since \( \chi_{\varepsilon,t} = ((\chi_{\varepsilon,t}^*)')^{-1} \) (cf. (A.4)), this concludes the proof of part (ii).

(iii). Let \( U((m,t), z) = V_t(z) - \tau_{\varepsilon,t}(m)z \). Then, using part (ii), Lemma A.1 and (B.18), we know that for \( k = 2,4 \) and for all \((m,t) \in K \times I \),
\[
 \int_{\mathbb{R}} \left| z - m \right|^k d\mu_{\varepsilon}^{\varepsilon,t,\chi_{\varepsilon,t}(m)}(z) = \int_{\mathbb{R}} \varepsilon^{-\frac{k}{2}} e^{-\frac{1}{2}U(m,t)(z)} \left( \frac{z - m}{s_t(m)} \right)^k e^{-\frac{1}{2}U(m,t)(z)} d\varepsilon \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}U(m,t)(z)} dz \quad \text{and} \quad t \rightarrow \frac{(k - 1)!}{V_t''((V_t')^{-1}(\tau_{\varepsilon,t}(m)))^\frac{k}{2}} + O_{K \times I} \left( \frac{1}{\varepsilon^{\frac{k-1}{2}}} \right). \tag{C.5}
\]

The dependence on \( \varepsilon \) of \( \tau_{\varepsilon,t} \) is of no problem here due to its uniform boundedness stated in part (ii). Then, for \( k = 2 \), the left-hand side of (C.5) equals \( s_t(m)^2 \) (cf. Lemma A.1). Thus, (C.5) proves the first claim in (C.2), since \((t,m,\varepsilon) \mapsto V_t''((V_t')^{-1}(\tau_{\varepsilon,t}(m))) \) is locally bounded. Moreover, due to Hölder’s inequality, to show the second claim in (C.2), it suffices to show that
\[
 \sup_{0 < \varepsilon < \varepsilon_0} \sup_{(m,t) \in K \times I} \int_{\mathbb{R}} \left| \frac{z - m}{s_t(m)} \right|^4 d\mu_{\varepsilon}^{\varepsilon,t,\chi_{\varepsilon,t}(m)}(z) < \infty. \tag{C.6}
\]
However, combining (C.5) for $k = 4$ and the first claim in (C.2), already implies (C.6). This conclude the proof of part (iii). □

Now we are in the position to prove the local Cramér theorem in the most general form that we use in this paper. The main ideas of the proof are the same as in [21, Proposition 31] or [25, Section 3]. The main difficulty here is to show that the estimates are uniform in $\varepsilon \ll 1$.

**Proposition C.2** Let $I, V$ and $\chi_{\varepsilon,t}$ be given as in Lemma C.1. Let $\chi_{N,\varepsilon,t} : \mathbb{R} \to \mathbb{R}$ be defined by

$$
\chi_{N,\varepsilon,t}(m) = -\frac{1}{N} \log \int_{P^{-1}(m)} e^{-\frac{1}{\varepsilon} \sum_{i=0}^{N-1} V_i(x_i)} dH_N^{-1}(x).
$$

(C.7)

Let $K \subset \mathbb{R}$ be compact. Then, for $\varepsilon$ small enough and for all $(m,t) \in K \times I$,

$$
e^{-N\chi_{N,\varepsilon,t}(m)} = e^{-N\chi_{\varepsilon,t}(m)} \sqrt{\chi''_{\varepsilon,t}(m)} \left( 1 + O_{K \times I} \left( \frac{1}{\sqrt{N}} \right) \right).
$$

(C.8)

**Proof.** Fix $(m,t) \in K \times I$. In this proof $C \in (0, \infty)$ denotes a constant, which is independent of $N$, $\varepsilon$, $m$ and $t$, and may change every time it appears.

Let $\chi^*_\varepsilon$, $s$ and $\mu^{\varepsilon,t,\sigma}$ be given as in Lemma C.1. In order to simplify the presentation here, for any function $f : \mathbb{R} \to \mathbb{R}$ and for all $z \in \mathbb{R}$, we abbreviate

$$
\langle f \rangle = \int_{\mathbb{R}} f(z) d\mu^{\varepsilon,t,\chi^*_\varepsilon}(z) \quad \text{and} \quad \hat{z} = \frac{z - m}{s_t(m)}.
$$

(C.9)

**Step 1.** [New representation of $e^{-N\chi_{N,\varepsilon,t}(m)}$.]

Let $(X_i)_i$ be a sequence of random variables that are independent and identically distributed with common law $\mu^{\varepsilon,t,\chi^*_\varepsilon}(m)$. Let

$$
\tilde{S}_{N,m} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} (X_i - m),
$$

(C.10)

and let $\tilde{g}_{N,m}$ denote the Lebesgue density of the distribution of $\tilde{S}_{N,m}$. As in [25, (31)], using the coarea formula, we have that

$$
\tilde{g}_{N,m}(0) = e^{N\chi_{\varepsilon,t}(m) - N\chi_{N,\varepsilon,t}(m)}.
$$

(C.11)

Moreover, let $g_{N,m}$ be the Lebesgue density of the distribution of

$$
S_{N,m} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} \frac{X_i - m}{st_{\varepsilon,t}(m)}.
$$

(C.12)

Then,

$$
g_{N,m}(0) = \tilde{g}_{N,m}(0) s_t(m) = \tilde{g}_{N,m}(0) \chi''_{\varepsilon,t}(m)^{-\frac{1}{2}}.
$$

(C.13)
Therefore, it suffices to show that for \( \varepsilon \) small enough,

\[
\left| g_{N,m}(0) - \frac{1}{\sqrt{2\pi}} \right| = O_{K \times I} \left( \frac{1}{\sqrt{N}} \right).
\]  

(C.14)

We show (C.14) by mimicking the arguments of the proof of [25, 3.1]. Therefore, as in [25, (44)], we apply the inverse Fourier transform to obtain that

\[
2\pi g_{N,m}(0) = \int_{\mathbb{R}} \left< e^{i \frac{1}{\sqrt{N}} \hat{\xi}} \right>^N d\hat{\xi},
\]  

(C.15)

and we split this integral according to some \( \delta < \hat{\delta} \), i.e.

\[
\int_{\mathbb{R}} \left< e^{i \frac{1}{\sqrt{N}} \hat{\xi}} \right>^N d\hat{\xi} = \int_{\{ |\hat{\xi}| \leq \delta \}} e^{-Nh\left( \frac{\hat{\xi}}{\sqrt{N}} \right)} d\hat{\xi} + \int_{\{ |\hat{\xi}| \geq \delta \}} \left< e^{i \frac{1}{\sqrt{N}} \hat{\xi}} \right>^N d\hat{\xi} =: I + II.
\]  

(C.16)

In the following we compute the asymptotic value of \( I \), and show that \( II \) is of lower order than \( I \).

**Step 2.** [Estimation of the term \( I \).] As in [25, (46)], we know that there exists \( \hat{\delta} > 0 \) and a complex-valued function \( h \) such that for all \( |\hat{\xi}| \leq \delta \),

\[
\left< e^{i\hat{\xi}} \right> = e^{-h(\hat{\xi})} \quad \text{and} \quad \left| h(\hat{\xi}) - \frac{1}{2}\hat{\xi}^2 \right| \leq C|\hat{\xi}|^3.
\]  

(C.17)

Indeed, this estimate follows from applying Taylor’s formula to the function \( \hat{\xi} \mapsto h(\hat{\xi}) \), and by using the fact that \( \sum_{k=1}^{3}|\hat{\xi}|^k \leq C \), which we have proven in Lemma C.1.

We argue similarly as in [25, (69)] to observe that (C.17) yields that (for \( \delta \) small enough)

\[
\text{Re} \left( Nh \left( \frac{\hat{\xi}}{\sqrt{N}} \right) \right) \geq \frac{|\hat{\xi}|^2}{4}.
\]  

(C.18)

Therefore, proceeding as in [25, p. 32], we infer that

\[
\left| e^{Nh\left( \frac{\hat{\xi}}{\sqrt{N}} \right)} - e^{-\frac{1}{2}\hat{\xi}^2} \right| \leq e^{-\frac{1}{2}\hat{\xi}^2}C|\hat{\xi}|^3 \sqrt{N},
\]  

(C.19)

which yields, as in [25, p. 32], to the estimate

\[
|I - \sqrt{2\pi}| = O_{K \times I} \left( \frac{1}{\sqrt{N}} \right).
\]  

(C.20)

**Step 3.** [Estimation of the term \( II \).] It remains to show that the term \( II \) is negligible. Applying the same arguments as in [25, p. 32], we see that it suffices to show that for all \( \delta < \hat{\delta} \) there exists \( \lambda < 1 \) such that

\[
\left| \left< e^{i\hat{\xi}} \right> \right| \leq \lambda \quad \text{for all} \quad |\hat{\xi}| \geq \delta.
\]  

(C.21)
And following the proof of [25, 3.4], we observe that (C.21) holds true once we show [25, (52)] and [25, (53)], i.e., if we show that

\[ \left\langle |\hat{\zeta}| \right\rangle \leq C \quad \text{and} \quad \left| \left\langle e^{i\hat{\zeta}} \right\rangle \right| \leq \frac{C}{|\hat{\zeta}|}. \tag{C.22} \]

The fact that \( \left\langle |\hat{\zeta}| \right\rangle \leq C \) was shown in Lemma C.1. To show the second claim in (C.22), let \( U_{(m,t)}(z) = V_t(z) - \tau_{t,t}(m)z \), where \( \tau_{t,t}(m) \) was introduced in Lemma C.1. Then, by partial integration (as in [25, p. 37]) and (C.2),

\[
\left| \left\langle e^{i\hat{\zeta}} \right\rangle \right| = \frac{s_t(m)}{|\hat{\zeta}|} \frac{1}{\varepsilon} \left| \int_{\mathbb{R}} e^{ix\hat{\zeta}} U'_{(m,t)}(z) e^{-\frac{1}{\varepsilon} U_{(m,t)}(z)} dz \right| \leq \frac{C}{|\hat{\zeta}|} \frac{1}{\sqrt{\varepsilon}} \int_{\mathbb{R}} \left| U'_{(m,t)}(z) \right| e^{-\frac{1}{\varepsilon} U_{(m,t)}(z)} dz.
\]

Let \( z_0 \) be the unique global minimum of \( U_{(m,t)} \), and let \( \rho = C'\sqrt{\varepsilon \log(\varepsilon^{-1})} \) for some \( C' > 0 \) large enough. Then, using the same arguments as in the proof of Lemma B.1, we see that the integral in the numerator in the right-hand side of (C.23) is concentrated around \( B_\rho(z_0) \), i.e.

\[
\int_{\mathbb{R}} |U'_{(m,t)}(z)| e^{-\frac{1}{\varepsilon} U_{(m,t)}(z)} dz = \int_{-\rho}^{\rho} |U'_{(m,t)}(z_0 + z)| e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0 + z)} dz + O_{Kx1}(e^2 e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0)}).
\]

Moreover, by Taylor’s formula for some \( \theta, \theta' \in [0, 1] \) (cf. (B.4)),

\[
\int_{-\rho}^{\rho} |U'_{(m,t)}(z_0 + z)| e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0 + z)} dz
= e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0)} \int_{-\rho}^{\rho} |z| U''_{(m,t)}(z_0 + \theta z)| e^{-\frac{1}{\varepsilon} U''_{(m,t)}(z_0 + \theta' z)} \frac{1}{\varepsilon^2} dz
\leq C e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0)} \int_{-\rho}^{\rho} |z| e^{-\frac{1}{\varepsilon} U''_{(m,t)}(z_0)} \frac{1}{\varepsilon^2} dz \leq C e^{-\frac{1}{\varepsilon} U_{(m,t)}(z_0)} \varepsilon.
\]

Combining (C.23), (C.24) and (C.25) and applying (B.1) to the denominator in the right-hand side of (C.23) yields the second claim in (C.22). This concludes the proof.

As a simple consequence of the ideas from the proof of Proposition C.2, we can state the result in a more precise way in the trivial case that \( V \) is a quadratic function and constant in \( t \). The result is given in the following lemma.

**Lemma C.3** Let \( V(z) = \frac{\alpha}{2} z^2 \) for some \( \alpha > 0 \). Let \( \chi_\varepsilon^*, \chi_\varepsilon \) and \( \mu_\varepsilon \chi_*^*(m) \) be given by (A.1), (A.2) and (A.3), respectively, with \( W \) replaced by \( \frac{1}{2} V \). Moreover, let \( \chi_{N,\varepsilon} \) be defined as in (C.7) but with \( V_t \) replaced by \( V \). Then, for all \( m \in \mathbb{R} \),

\[ e^{-N\chi_{N,\varepsilon}(m)} = e^{-N\chi_\varepsilon(m)} \frac{\sqrt{\chi_*^\prime(m)}}{\sqrt{2\pi}} \tag{C.26} \]

**Proof.** Note that by a simple computation, for all \( \sigma, m \in \mathbb{R} \),

\[ \chi_\varepsilon(m) = \frac{\alpha}{2\varepsilon} m^2 - \frac{1}{2} \log \left( 2\pi \frac{\varepsilon}{\alpha} \right) \quad \text{and} \quad \mu_\varepsilon \chi_*^*(m) = e^{-\frac{\varepsilon}{2}(\varepsilon-m)^2} \frac{1}{\sqrt{2\pi}} dz. \tag{C.27} \]
Using the same notation and the same arguments as in Step 1 of the proof of Proposition C.2, we see that it suffices to show that

\[ g_{N,m}(0) = \frac{1}{\sqrt{2\pi}}. \]  

(C.28)

However, since \( \mu_{\varepsilon, \chi^{(m)}}(z) \) is a Gaussian measure, the claim (C.28) is a simple consequence of the stability of Gaussian measures under convolution. This concludes the proof. \( \square \)

Acknowledgement. The authors would like to give many thanks to Anton Bovier, Dmitry Ioffe, Lorenzo Dello Schiavo and André Schlichting for many useful discussions and suggestions.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] F. Barret. Sharp asymptotics of metastable transition times for one dimensional SPDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):129–166, 2015.

[3] F. Barret, A. Bovier, and S. Mélaéard. Uniform estimates for metastable transition times in a coupled bistable system. *Electron. J. Probab.*, 15:no. 12, 323–345, 2010.

[4] K. Bashiri and A. Bovier. Gradient flow approach to local mean-field spin systems. *Stochastic Processes and their Applications*, https://doi.org/10.1016/j.spa.2019.05.006, 2019.

[5] N. Berglund. Kramers' law: validity, derivations and generalisations. *Markov Process. Related Fields*, 19(3):459–490, 2013.

[6] N. Berglund, G. Di Gesù, and H. Weber. An Eyring-Kramers law for the stochastic Allen-Cahn equation in dimension two. *Electron. J. Probab.*, 22:Paper No. 41, 27, 2017.

[7] N. Berglund, B. Fernandez, and B. Gentz. Metastability in interacting nonlinear stochastic differential equations. II. Large-\(N\) behaviour. *Nonlinearity*, 20(11):2583–2614, 2007.

[8] N. Berglund and B. Gentz. Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond. *Electron. J. Probab.*, 18:no. 24, 58, 2013.

[9] J. Bewersdorff. *Algebra für Einsteiger*. Friedr. Vieweg & Sohn, Wiesbaden, second edition, 2004. Von der Gleichungsauflösung zur Galois-Theorie.

[10] A. Bovier. *Gaussian processes on trees*, volume 163 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017. From spin glasses to branching Brownian motion.

[11] A. Bovier and F. den Hollander. *Metastability*, volume 351 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2015. A potential-theoretic approach.

[12] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theory Related Fields*, 119(1):99–161, 2001.
[13] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low lying spectra in reversible Markov chains. Comm. Math. Phys., 228(2):219–255, 2002.

[14] 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. (JEMS), 6(4):399–424, 2004.

[15] M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares. Metastable behavior of stochastic dynamics: a pathwise approach. J. Statist. Phys., 35(5-6):603–634, 1984.

[16] D. A. Dawson and J. Gärtner. Large deviations and tunnelling for particle systems with mean field interaction. C. R. Math. Rep. Acad. Sci. Canada, 8(6):387–392, 1986.

[17] D. A. Dawson and J. Gärtner. Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. Mem. Amer. Math. Soc., 78(398):iv+94, 1989.

[18] A. Dembo and O. Zeitouni. Large deviations techniques and applications, volume 38 of Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[19] R. S. Ellis, J. L. Monroe, and C. M. Newman. The ghs and other correlation inequalities for a class of even ferromagnets. Comm. Math. Phys., 46(2):167–182, 1976.

[20] L. C. Evans and R. F. Gariepy. Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[21] N. Grunewald, F. Otto, C. Villani, and M. G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. Ann. Inst. Henri Poincaré Probab. Stat., 45(2):302–351, 2009.

[22] S. Herrmann and J. Tugaut. Non-uniqueness of stationary measures for self-stabilizing processes. Stochastic Process. Appl., 120(7):1215–1246, 2010.

[23] Y. Kwon and G. Menz. Decay of correlations and uniqueness of the infinite-volume Gibbs measure of the canonical ensemble of 1d-lattice systems. J. Stat. Phys., 176(4):836–872, 2019.

[24] C. Landim. Metastable Markov chains. Probab. Surv., 16:143–227, 2019.

[25] G. Menz and F. Otto. Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. Ann. Probab., 41(3B):2182–2224, 2013.

[26] P. E. Müller. Limiting properties of a continuous local mean-field interacting spin system: Hydrodynamic limit, propagation of chaos, energy landscape and large deviations. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2016.

[27] E. Olivieri and M. E. Vares. Large deviations and metastability, volume 100 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2005.

[28] F. Otto. The geometry of dissipative evolution equations: The porous medium equation. Communications in Partial Differential Equations, 26(1-2):101–174, 2001.

[29] I. Seo. Analysis of metastable behavior via solutions of poisson equations. Preprint. arXiv:1905.00743, 2019.

[30] M.-K. von Renesse and K.-T. Sturm. Entropic measure and Wasserstein diffusion. Ann. Probab., 37(3):1114–1191, 2009.