

BROWNIAN OCCUPATION MEASURES, COMPACTNESS AND LARGE DEVIATIONS

BY CHIRANJIB MUKHERJEE[†] AND S. R. S. VARADHAN^{‡,1}

*Courant Institute New York[†] and WIAS Berlin,[†]
and Courant Institute New York[‡]*

In proving large deviation estimates, the lower bound for open sets and upper bound for compact sets are essentially local estimates. On the other hand, the upper bound for closed sets is global and compactness of space or an exponential tightness estimate is needed to establish it. In dealing with the occupation measure $L_t(A) = \frac{1}{t} \int_0^t \mathbb{1}_A(W_s) ds$ of the d -dimensional Brownian motion, which is not positive recurrent, there is no possibility of exponential tightness. The space of probability distributions $\mathcal{M}_1(\mathbb{R}^d)$ can be compactified by replacing the usual topology of weak convergence by the vague topology, where the space is treated as the dual of continuous functions with compact support. This is essentially the one point compactification of $\mathcal{M}_1(\mathbb{R}^d)$ by adding a point at ∞ that results in the compactification of $\mathcal{M}_1(\mathbb{R}^d)$ by allowing some mass to escape to the point at ∞ . If one were to use only test functions that are continuous and vanish at ∞ , then the compactification results in the space of sub-probability distributions $\mathcal{M}_{\leq 1}(\mathbb{R}^d)$ by ignoring the mass at ∞ .

The main drawback of this compactification is that it ignores the underlying translation invariance. More explicitly, we may be interested in the space of equivalence classes of orbits $\widetilde{\mathcal{M}}_1 = \widetilde{\mathcal{M}}_1(\mathbb{R}^d)$ under the action of the translation group \mathbb{R}^d on $\mathcal{M}_1(\mathbb{R}^d)$. There are problems for which it is natural to compactify this space of orbits. We will provide such a compactification, prove a large deviation principle there and give an application to a relevant problem.

1. Motivation and introduction.

1.1. *Motivation.* We start with the Wiener measure \mathbb{P} on $\Omega = C_0([0, \infty); \mathbb{R}^d)$ corresponding to the d -dimensional Brownian motion $W = (W_t)_{t \geq 0}$ starting from the origin. Our result is motivated by the following set up. Let L_t denote the normalized occupation measure of the Brownian motion until time t , that is,

$$(1) \quad L_t = \frac{1}{t} \int_0^t ds \delta_{W_s}.$$

Received April 2014; revised September 2015.

¹Supported in part by NSF Grant DMS-12-08334.

MSC2010 subject classifications. Primary 60J65, 60J35; secondary 60F10.

Key words and phrases. Occupation measures, large deviations, compactification, polaron problem.

This is a random element of $\mathcal{M}_1 = \mathcal{M}_1(\mathbb{R}^d)$, the space of probability measures on \mathbb{R}^d . We are interested in the transformed measure

$$(2) \quad \widehat{\mathbb{P}}_t(A) = \frac{1}{Z_t} \mathbb{E}\{\mathbb{1}_A \exp\{tH(L_t)\}\}$$

with A being a measurable set in the path space of the Brownian motion and

$$(3) \quad H(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} V(x - y)\mu(dx)\mu(dy).$$

Here, $V(\cdot)$ is a continuous function on \mathbb{R}^d vanishing at infinity and

$$Z_t = \mathbb{E}\{\exp\{tH(L_t)\}\}$$

is the normalizing constant or the *partition function*. For $d = 3$ and $V(x) = \frac{1}{|x|}$, it is known (see [4]) that,

$$(4) \quad \begin{aligned} & \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}\{\exp\{tH(L_t)\}\} \\ &= \sup_{\substack{\psi \in H^1(\mathbb{R}^d) \\ \|\psi\|_2=1}} \left\{ \iint_{\mathbb{R}^d \times \mathbb{R}^d} dx dy V(x - y)\psi^2(x)\psi^2(y) - \frac{1}{2} \|\nabla \psi\|_2^2 \right\}, \end{aligned}$$

where $H^1(\mathbb{R}^d)$ is the usual Sobolev space of square integrable functions in with their gradient in $L^2(\mathbb{R}^d)$. For $d = 3$, $V(x) = \frac{1}{|x|}$, this variational formula has also been analyzed by Lieb (see [5]) who proved that there is a maximizer which is unique except for spatial translations. In other words, if \mathfrak{m} denotes the set of maximizing densities, then

$$(5) \quad \mathfrak{m} = \{\mu_0 \star \delta_x : x \in \mathbb{R}^3\},$$

where μ_0 is a probability measure with a density ψ_0^2 so that ψ_0 maximizes the variational problem (4).

Given (4) and (5), we expect that the asymptotic distribution of L_t under $\widehat{\mathbb{P}}_t$ to be concentrated around \mathfrak{m} . Indeed, we would like to show that for very $\varepsilon > 0$,

$$(6) \quad \lim_{t \rightarrow \infty} \widehat{\mathbb{P}}_t\{L_t \notin U_\varepsilon(\mathfrak{m})\} = 0,$$

where $U_\varepsilon(\mathfrak{m})$ is a (weak) neighborhood of \mathfrak{m} . In fact, we can write

$$\begin{aligned} \widehat{\mathbb{P}}_t\{L_t \notin U_\varepsilon(\mathfrak{m})\} &= \frac{\mathbb{E}[\mathbb{1}_{L_t \notin U_\varepsilon(\mathfrak{m})} \exp(tH(L_t))]}{\mathbb{E}[\exp(tH(L_t))]} \\ &= \frac{\mathbb{E}[\exp(tF(L_t))]}{\mathbb{E}[\exp(tH(L_t))]}, \end{aligned}$$

where

$$(7) \quad F(\mu) = \begin{cases} H(\mu) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} V(x - y)\mu(dx)\mu(dy), & \text{if } \mu \notin U_\varepsilon(\mathfrak{m}), \\ -\infty, & \text{else.} \end{cases}$$

Let us pretend that we have a *strong* Donsker–Varadhan large deviation principle for L_t in $\mathcal{M}_1(\mathbb{R}^d)$ under the weak topology; see (29)–(30) and the following remarks for a precise definition. Then, using Varadhan’s lemma [ignoring the lack of upper semicontinuity of F coming from the singularity for $V(x) = 1/|x|$] we could (formally) conclude that $\widehat{\mathbb{P}}_t(L_t \notin U_\varepsilon(m))$ decays exponentially fast in t .

However, the lack of a strong large deviation principle turns out to be crucial issue. To circumvent this problem, the space $\mathcal{M}_1(\mathbb{R}^d)$ has to be “compactified.” This can be done by replacing the usual topology of weak convergence by the “vague topology,” where the space is treated as the dual of continuous functions with compact support. This is essentially the one point compactification of \mathbb{R}^d by adding a point at ∞ that results in the compactification of $\mathcal{M}_1(\mathbb{R}^d)$ by allowing some mass to escape to the point at ∞ . If one were to use only test functions that are continuous and vanish at ∞ , then the compactification results in the space of sub-probability distributions $\mathcal{M}_{\leq 1}(\mathbb{R}^d)$ by ignoring the mass at ∞ .

Let us also mention that, for (4), in [4], the lack of compactness of the state space was handled by replacing Brownian motion by Ornstein–Uhlenbeck (O-U) process on \mathbb{R}^d whose occupation measure, unlike Brownian motion, satisfies a strong large deviation principle. Exploiting the positive definiteness of $V(x) = \frac{1}{|x|}$ the authors show that the total mass $\mathbb{E}\{\exp\{tH(L_t)\}\}$ is dominated by the same expectation with respect to the Ornstein–Uhlenbeck process. This monotonicity combined with strong large deviations for the O-U process proves (4). However, no such monotonicity is available to us in the complement of the neighborhood of m (i.e., for the term $\mathbb{E}[\mathbb{1}_{L_t \notin U_\varepsilon(m)} \exp(tH(L_t))]$). Another possibility is to replace \mathbb{R}^d by a large torus and “fold” L_t in the torus and use a similar monotonicity of the total masses (see [1, 3]). Although these methods work well for deriving asymptotic behavior of the partition function, questions on the path measures $\widehat{\mathbb{P}}_t$ can not be handled so well in this manner. In particular, these methods ignore the underlying *translation invariance* of some relevant models from statistical mechanics, models which depend on shift-invariant functionals of the occupation measures L_t , like the functional $H(\mu) = H(\mu \star \delta_x)$ for all $x \in \mathbb{R}^d$, defined in (3). Motivated by this, we naturally consider the quotient space

$$\widetilde{\mathcal{M}}_1(\mathbb{R}^d) = \mathcal{M}_1(\mathbb{R}^d) / \sim$$

under spatial shifts and are led to a robust theory of compactification of this space. Let us briefly sketch the main idea here.

1.2. *Translation invariant compactification: The central idea.* Note that $\mathcal{M}_1(\mathbb{R}^d)$ fails to be compact in the weak topology for several reasons. For instance, if we take a Gaussian with a very large variance, the mass can spread very thin and totally disintegrate into dust. Also, a mixture like $\frac{1}{2}(\mu \star \delta_{a_n} + \mu \star \delta_{-a_n})$ splits into two (or more) widely separated pieces as $a_n \rightarrow \infty$. To compactify this space we should be allowed to “center” each piece separately as well as to allow some mass to be “thinly spread and disappear.”

The intuitive idea, starting with a sequence of probability distributions $(\mu_n)_n$ in \mathbb{R}^d is to identify a compact region where μ_n has its largest accumulation of mass. This is given by its *concentration function* defined by

$$q_n(r) = \sup_{x \in \mathbb{R}^d} \mu_n(B_r(x)).$$

By choosing subsequences, we can assume that $q_n(r) \rightarrow q(r)$ as $n \rightarrow \infty$ and $q(r) \rightarrow p_1 \in [0, 1]$ as $r \rightarrow \infty$. Then there is a shift $\lambda_n = \mu_n \star \delta_{a_n}$ which converges along a subsequence vaguely to a sub-probability measure α_1 of mass p_1 . This means λ_n can be written as $\alpha_n + \beta_n$ so that $\alpha_n \Rightarrow \alpha_1$ weakly and we recover the partial mass $p_1 \in [0, 1]$. We peel off α_n from λ_n and repeat the same process for β_n to get convergence along a further subsequence. We go on recursively to get convergence of one component at a time along further subsequences in the space of sub-probability measures, *modulo spatial shifts*. The picture is, μ_n roughly concentrates on *widely separated* compact pieces of masses $\{p_j\}_{j \in \mathbb{N}}$ while the rest of the mass $1 - \sum_j p_j$ leaks out.

In other words, given any sequence $\tilde{\mu}_n$ of equivalence classes in $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$, which is the quotient space of $\mathcal{M}_1(\mathbb{R}^d)$ under spatial shifts, there is a subsequence which converges (in a sense which we do not make precise yet) to an element $\{\tilde{\alpha}_1, \tilde{\alpha}_2, \dots\}$, a collection of equivalence classes of sub-probabilities α_j of masses $0 \leq p_j \leq 1$, $j \in \mathbb{N}$.² The space of such collections of equivalence classes is the compactification of $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$ and in this space we are able to prove a strong large deviation principle for the distribution of the equivalence classes \tilde{L}_t of L_t . This, combined with the shift invariant structure of $V(x - y)$, enables us to prove (6).

Finally, although we were motivated by asymptotic study of path measures of mean-field type interactions for Brownian motion, it can also be applied to study a wider class of problems that involve translation invariant functionals of processes with independent increments.

Let us describe the organization of the rest of the article. In Section 2, we collect some basic facts about weak and vague convergence, introduce a class of relevant test functions and characterize notions of *total disintegration of measures* as well as measures being *widely separated* in terms of test integrals with respect to the corresponding test functions. In Section 3, we introduce a space $\tilde{\mathcal{X}}$ and a metric \mathbf{D} giving rise a notion of topology and convergence in this space. Here we also prove that $\tilde{\mathcal{X}}$ is the desired compactification of the quotient space $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$. Section 4 is devoted to proving a strong large deviation principle for the distribution of the equivalence class \tilde{L}_t in $\tilde{\mathcal{X}}$ and in Section 5 we provide the application to the asymptotics of the path measures $\hat{\mathbb{P}}_t(\tilde{L}_t \in \cdot)$.

²For example, let μ_n be a sequence which is a mixture of three Gaussians, one with mean 0 and variance 1, one with mean n and variance 1 and one with mean 0 and variance n , each with equal weight $\frac{1}{3}$. Then the limiting object is the collection $\{\tilde{\alpha}_1, \tilde{\alpha}_1\}$, where $\tilde{\alpha}_1$ is the equivalence class of a Gaussian with variance 1 and weight $\frac{1}{3}$.

2. Topologies on measures. We denote by $\mathcal{M}_1 = \mathcal{M}_1(\mathbb{R}^d)$ the space of probability distributions on \mathbb{R}^d and by $\tilde{\mathcal{M}}_1 = \mathcal{M}_1 / \sim$ the quotient space of \mathcal{M}_1 under the action of \mathbb{R}^d (as an additive group on \mathcal{M}_1). For any $\mu \in \mathcal{M}_1$, its orbit is defined by $\tilde{\mu} = \{\mu \star \delta_x : x \in \mathbb{R}^d\} \in \tilde{\mathcal{M}}_1$.

2.1. *The weak and the vague topology.* We turn to two natural topologies on \mathcal{M}_1 . In the *weak topology*, a sequence μ_n in \mathcal{M}_1 converges to μ , denoted by $\mu_n \Rightarrow \mu$, if

$$(8) \quad \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f(x) \mu_n(dx) = \int_{\mathbb{R}^d} f(x) \mu(dx),$$

for all bounded continuous functions on \mathbb{R}^d . On the other hand, in the *vague topology* for the convergence of μ_n to μ , denoted by $\mu_n \hookrightarrow \mu$, we only require (8) for continuous functions with compact support. It continues to hold for continuous functions that tend to 0 as $|x| \rightarrow \infty$. Note that the total mass of probability measures, which is conserved in the weak convergence, is not necessarily conserved under vague convergence—a salient feature which distinguishes these two topologies. If we denote by $\mathcal{M}_{\leq 1} = \mathcal{M}_{\leq 1}(\mathbb{R}^d)$ the space of all sub-probability measures (non-negative measures with total mass less than or equal to one), then both topologies carry over to $\mathcal{M}_{\leq 1}$ with the same requirements.

We collect some standard facts as a lemma which will be relevant for us.

- LEMMA 2.1. (i) If $\mu_n \hookrightarrow \mu$ in $\mathcal{M}_{\leq 1}$, then $\mu(\mathbb{R}^d) \leq \liminf_{n \rightarrow \infty} \mu_n(\mathbb{R}^d)$.
 (ii) If $\mu_n \hookrightarrow \mu$ in $\mathcal{M}_{\leq 1}$ and $\mu_n(\mathbb{R}^d) \rightarrow \mu(\mathbb{R}^d)$, then $\mu_n \Rightarrow \mu$ in $\mathcal{M}_{\leq 1}$.
 (iii) While \mathcal{M}_1 is a closed subset of the space $\mathcal{M}_{\leq 1}$ in the weak topology, it is dense in $\mathcal{M}_{\leq 1}$ in the vague topology.
 (iv) The space $\mathcal{M}_{\leq 1}$ is compact in the vague topology.

We will also need the following elementary lemma.

LEMMA 2.2. If $\mu_n \hookrightarrow \alpha$ in $\mathcal{M}_{\leq 1}$, then μ_n can be written as $\mu_n = \alpha_n + \beta_n$ where $\alpha_n \Rightarrow \alpha$ and $\beta_n \hookrightarrow 0$.

PROOF. We will denote by $B(x, r)$ the ball of radius $r > 0$ around the point $x \in \mathbb{R}^d$. If $\mu_n \hookrightarrow \alpha$ then

$$\lim_{n \rightarrow \infty} \mu_n(B(0, r)) = \alpha(B(0, r)),$$

for all but at most countably values of r and $\alpha(\mathbb{R}^d)$ can be recovered as the limit

$$\alpha(\mathbb{R}^d) = \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} \mu_n(B(0, r)).$$

Hence, given any $r > 0$, there is $n_r \in \mathbb{N}$ such that for $n \geq n_r$ we have,

$$\mu_n(B(0, r)] \leq \alpha(\mathbb{R}^d) + \frac{1}{r}.$$

Without loss of generality, we can assume that n_r is nondecreasing with r . If we define

$$R_n = \sup\{r > 0 : n_r \leq n\},$$

then $R_n \rightarrow \infty$ and

$$\mu_n(B(0, R_n)) \leq \alpha(\mathbb{R}^d) + \frac{1}{R_n}.$$

If we take α_n and β_n as the restrictions of μ_n to $B(0, R_n)$ and $B(0, R_n)^c$ respectively, $\alpha_n \hookrightarrow \alpha$ and $\alpha_n(\mathbb{R}^d) \rightarrow \alpha(\mathbb{R}^d)$. Therefore, by Lemma 2.1 part (ii), $\alpha_n \Rightarrow \alpha$. Furthermore, for any given $r > 0$, eventually $\beta_n(B(0, r)) = 0$ and hence $\beta_n \hookrightarrow 0$. □

2.2. *The space \mathcal{F} of test functions.* For our desired compactification, we need to develop a suitable topology on the quotient space $\widetilde{\mathcal{M}}_1$ via convergence of test integrals. For this, we first need to characterize a suitable class of continuous functions (or rather, functionals) on $\widetilde{\mathcal{M}}_1$.

We fix a positive integer $k \geq 2$. Let \mathcal{F}_k be the space of continuous functions $f : (\mathbb{R}^d)^k \rightarrow \mathbb{R}$ that are *translation invariant*, i.e.,

$$f(x_1 + y, \dots, x_k + y) = f(x_1, \dots, x_k) \quad \forall y, x_1, \dots, x_k \in \mathbb{R}^d$$

and *vanish at infinity*, in the sense,

$$\lim_{\max_{i \neq j} |x_i - x_j| \rightarrow \infty} f(x_1, \dots, x_k) = 0.$$

In other words, $f(x_1, \dots, x_k)$ depends only on the differences $\{x_i - x_j\}_{i,j}$. A typical example of a function $f \in \mathcal{F}_2$ could be $f(x_1, x_2) = V(x_1 - x_2)$ where $V(\cdot)$ is a continuous function such that $V(x) \rightarrow 0$ as $|x| \rightarrow \infty$. Note that, each $f \in \mathcal{F}_k$ can be interpreted as a continuous function of $k - 1$ variables vanishing at infinity. Hence, for each $k \geq 2$, \mathcal{F}_k is a separable space under the uniform metric. Hence, if we denote by

$$\mathcal{F} = \bigcup_{k \geq 2} \mathcal{F}_k,$$

then we can choose a countable dense subset for each \mathcal{F}_k and ordering all of them as a single countable sequence $\{f_r(x_1, \dots, x_{k_r}) : r \in \mathbb{N}\}$ we obtain a countable dense subset of \mathcal{F} .

For any $\mu \in \mathcal{M}_1$ and $f \in \mathcal{F}$, we define the function

$$\Lambda(f, \mu) = \int_{(\mathbb{R}^d)^k} f(x_1, \dots, x_k) \mu(dx_1) \cdots \mu(dx_k).$$

Note that, because of translation invariance of f , $\Lambda(f, \mu)$ depends only on the orbit $\tilde{\mu} \in \widetilde{\mathcal{M}}_1$ for any fixed $f \in \mathcal{F}$. As it will turn out, these are natural continuous

functions to consider on $\widetilde{\mathcal{M}}_1$. Given any sequence $(\mu_n)_n$ in \mathcal{M}_1 , because there is a countable dense set $\{f_r\}$, by diagonalization one can choose a subsequence such that along the subsequence (denoted again by μ_n), the limit

$$\Lambda(f) = \lim_{n \rightarrow \infty} \Lambda(f, \mu_n),$$

exists. To compactify the space $\widetilde{\mathcal{M}}_1$ we will determine what the set of possible limits are; see Section 3.

2.3. *Total disintegration of a sequence of measures.* We say that a sequence (μ_n) in $\mathcal{M}_{\leq 1}$ *totally disintegrates* if for any positive $r < \infty$,

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \mu_n(B(x, r)) = 0.$$

A typical example of a totally disintegrating sequence μ_n of measures is a centered Gaussian with covariance matrix $n\mathbf{Id}$.

The following facts determine equivalent criteria for total disintegration of a sequence of measures and it is useful to collect them.

LEMMA 2.3. *Let $(\mu_n)_n$ be a sequence in $\mathcal{M}_{\leq 1}$. The following facts are equivalent.*

(a) *There exists a continuous function $V(x) > 0$ on \mathbb{R}^d , with $\lim_{|x| \rightarrow \infty} V(x) = 0$, such that*

(9)
$$\lim_{n \rightarrow \infty} \iint_{\mathbb{R}^{2d}} V(x - y) \mu_n(dx) \mu_n(dy) = 0.$$

(b)

(10)
$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \mu_n(B(x, r)) = 0.$$

(c) *For any continuous function $V(x)$ with $\lim_{|x| \rightarrow \infty} V(x) = 0$,*

(11)
$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \int V(x - y) \mu_n(dy) = 0.$$

(d) *For any continuous function $V(x)$ with $\lim_{|x| \rightarrow \infty} V(x) = 0$,*

(12)
$$\lim_{n \rightarrow \infty} \iint_{\mathbb{R}^{2d}} V(x - y) \mu_n(dx) \mu_n(dy) = 0.$$

Let the sequence μ_n of measures satisfy any of the above. Then for any $k \geq 2$ and $f \in \mathcal{F}_k$

(13)
$$\lim_{n \rightarrow \infty} \int \cdots \int_{\mathbb{R}^{dk}} f(x_1, \dots, x_k) \mu_n(dx_1) \cdots \mu_n(dx_k) = 0.$$

PROOF. (a) \Rightarrow (b). Let $r > 0$ be given. Since $V(x) > 0$ and continuous, there exists $\delta > 0$ such that $V(x) \geq \delta$ on $B(0, 2r)$. Then,

$$\begin{aligned} \iint_{\mathbb{R}^{2d}} V(x - y)\mu_n(dx)\mu_n(dy) &\geq \delta \int_{|x-y|\leq 2r} \mu_n(dx)\mu_n(dy) \\ &\geq \delta \sup_{x \in \mathbb{R}^d} \{\mu_n(B(x, r))\}^2. \end{aligned}$$

(b) \Rightarrow (c). Let $\varepsilon_M = \sup_{|x|\geq M} |V(x)|$. Then $\lim_{M \rightarrow \infty} \varepsilon_M = 0$ and

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} \int |V(x - y)|\mu_n(dy) &\leq \sup_{x \in \mathbb{R}^d} \int_{B(x, M)} |V(x - y)|\mu_n(dy) \\ &\quad + \sup_{x \in \mathbb{R}^d} \int_{B(x, M)^c} |V(x - y)|\mu_n(dy) \\ &\leq \|V\|_\infty \sup_{x \in \mathbb{R}^d} \mu_n[B(x, M)] + \varepsilon_M. \end{aligned}$$

Therefore,

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \int |V(x - y)|\mu_n(dy) \leq \varepsilon_M,$$

for any M . Since $\varepsilon_M \rightarrow 0$ if we let $M \rightarrow \infty$, we get the claim.

(c) \Rightarrow (d). Observe that, since $\mu_n(\mathbb{R}^d) \leq 1$,

$$\iint_{\mathbb{R}^{2d}} V(x - y)\mu_n(dx)\mu_n(dy) \leq \sup_{x \in \mathbb{R}^d} \int V(x - y)\mu_n(dy).$$

(d) \Rightarrow (a). This is obvious.

For the last part, for $k > 2$ we define

$$W(x_1, x_2) = \sup_{x_3, \dots, x_k} |f(x_1, \dots, x_k)|.$$

Note that $W \in \mathcal{F}_2$ and so it is of the form $V(x_1 - x_2)$. Since

$$\begin{aligned} \int \cdots \int_{\mathbb{R}^{dk}} |f(x_1, \dots, x_k)|\mu_n(dx_1) \cdots \mu_n(dx_k) &\leq \int \int_{\mathbb{R}^{2d}} W(x_1, x_2)\mu_n(dx_1)\mu_n(dx_2) \\ &= \int \int_{\mathbb{R}^{2d}} V(x - y)\mu_n(dx)\mu_n(dy), \end{aligned}$$

the lemma is proved. \square

2.4. *Widely separated sequences of measures.* We now need a working definition of two *widely separated* sequence of measures. We say that two sequences $(\alpha_n)_n$ and $(\beta_n)_n$ in $\mathcal{M}_{\leq 1}$ are widely separated, if for some strictly positive function V on \mathbb{R}^d which is continuous and vanishes at infinity,

$$(14) \quad \lim_{n \rightarrow \infty} \int V(x - y)\alpha_n(dx)\beta_n(dy) = 0.$$

Note that if a sequence $(\mu_n)_n$ in $\mathcal{M}_{\leq 1}$ satisfies (9), then, because of (11), it is widely separated from any arbitrary sequence of measures in $\mathcal{M}_{\leq 1}$.

LEMMA 2.4. *Let $(\alpha_n)_n$ and $(\beta_n)_n$ be two widely separated sequences in $\mathcal{M}_{\leq 1}$. Then:*

(i) *For any continuous function W in \mathbb{R}^d vanishing at infinity*

$$\lim_{n \rightarrow \infty} \int W(x - y) \alpha_n(dx) \beta_n(dy) = 0.$$

(ii) *For every $k \geq 2$ and $f \in \mathcal{F}_k$,*

$$(15) \quad \lim_{n \rightarrow \infty} \left| \int f(x_1, \dots, x_k) \prod_{i=1}^k [\alpha_n + \beta_n](dx_i) - \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha_n(dx_i) - \int f(x_1, \dots, x_k) \prod_{i=1}^k \beta_n(dx_i) \right| = 0.$$

PROOF. Let W be any continuous function \mathbb{R}^d vanishing at infinity. Since $(\alpha_n)_n$ and $(\beta_n)_n$ are widely separated, for some strictly positive V which is continuous and vanishes at infinity,

$$\lim_{n \rightarrow \infty} \int V(x - y) \alpha_n(dx) \beta_n(dy) = 0.$$

Furthermore, given any $\varepsilon > 0$, there is a constant $C_\varepsilon > 0$ such that

$$|W(x)| \leq C_\varepsilon V(x) + \varepsilon.$$

Then

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \int |W(x - y)| \alpha_n(dx) \beta_n(dy) \\ & \leq C_\varepsilon \limsup_{n \rightarrow \infty} \int V(x - y) \alpha_n(dx) \beta_n(dy) + \varepsilon = \varepsilon. \end{aligned}$$

This proves the first part (i).

For the second part (ii), if we take $k = 2$ and expand the product

$$\prod_{i=1}^2 (\alpha_n + \beta_n)(dx_i),$$

it is seen that all the cross terms are controlled by (14) and are negligible, by the first part (i), as $f(x_1, x_2) = W(x_1 - x_2)$ for some continuous W vanishing at infinity. The general case $k \geq 3$ follows easily. \square

3. Compactification of $\tilde{\mathcal{M}}_1$: The space $\tilde{\mathcal{X}}$. We turn to the central issue of \mathcal{M}_1 failing to be compact in the weak topology. As mentioned before some typical reasons for this could be as follows: The location of the mass can shift away to ∞ as in $\mu_n = \mu * \delta_{a_n}$ with $a_n \rightarrow \infty$, or it can split into two (or more) pieces like in $\mu_n = \frac{1}{2}[\mu * \delta_{a_n} + \mu * \delta_{-a_n}]$, or it can also totally disintegrate into dust like a Gaussian with a large variance. One imagines, in the limit, an empty, finite or countable collection I of mass distributions $\{\alpha_i: i \in I\}$ that are widely separated with total mass $\sum_{i \in I} \alpha_i(\mathbb{R}^d) = p \leq 1$ and the remaining mass $1 - p$ having totally disintegrated. Therefore, a natural “compactification” could be a space $\tilde{\mathcal{X}}$ of empty, finite or countable collections of orbits $\{\tilde{\alpha}_i: i \in I\}$ of sub-probability distributions α_i having masses p_i with $p = \sum_i p_i \leq 1$.

3.1. *The space $\tilde{\mathcal{X}}$ and a metric \mathbf{D} .* We define

$$(16) \quad \tilde{\mathcal{X}} = \left\{ \xi: \xi = \{\tilde{\alpha}_i\}_{i \in I}, \alpha_i \in \mathcal{M}_{\leq 1}, \sum_{i \in I} \alpha_i(\mathbb{R}^d) \leq 1 \right\}.$$

We make some remarks about the above definition.

REMARK 1. First note that, in order to keep notation short, we suppressed the fact that the index set I above ranges over empty, finite or countably many collections. Furthermore, we will write any typical element $\xi \in \tilde{\mathcal{X}}$ as $\xi = \{\alpha_i\}$ with the understanding that either the collection is empty or i ranges over a finite or countable set.

REMARK 2. Note that any element $\alpha \in \mathcal{M}_{\leq 1}$ in the orbit $\tilde{\alpha}$ has the same total mass $\alpha(\mathbb{R}^d)$. Hence, for any element $\xi = \{\tilde{\alpha}_i\} \in \tilde{\mathcal{X}}$, $p_i = \alpha_i(\mathbb{R}^d)$ will denote the total mass of any candidate α_i in the orbit $\tilde{\alpha}_i$ and $\sum_i p_i = p \leq 1$. If the collection is empty, then $p = 0$ vacuously.

REMARK 3. Note that in any element $\xi = \{\tilde{\alpha}_i\}$ of $\tilde{\mathcal{X}}$, an orbit α_i could be repeated more than once. We call the number of occurrences of an orbit in an element ξ its *multiplicity*.

Recall the class of functions \mathcal{F}_k for $k \geq 2$ and note that $\mathcal{F} = \bigcup_{k \geq 2} \mathcal{F}_k$. We now introduce a metric on $\tilde{\mathcal{X}}$ with the following convergence criterion. We want a sequence $(\xi_n)_n$ to converge to ξ in the space $\tilde{\mathcal{X}}$ under the desired metric, if the sequence

$$\Lambda(f, \xi_n) = \sum_{\tilde{\alpha}_n \in \xi_n} \int f(x_1, \dots, x_k) \alpha_n(dx_1) \cdots \alpha_n(dx_k)$$

converges to the corresponding expression

$$\Lambda(f, \xi) = \sum_{\tilde{\alpha} \in \xi} \int f(x_1, \dots, x_k) \alpha(dx_1) \cdots \alpha(dx_k)$$

for every $f \in \mathcal{F}$. Recall that the value of $\int f(x_1, \dots, x_k) \alpha(dx_1) \cdots \alpha(dx_k)$ depends only on the orbit $\tilde{\alpha}$ since f is translation invariant. We also remark that if ξ is empty then $\Lambda(f, \xi) = 0$ for all $f \in \mathcal{F}$.

For any $\xi_1, \xi_2 \in \tilde{\mathcal{X}}$, we define

$$(17) \quad \mathbf{D}(\xi_1, \xi_2) = \sum_{r=1}^{\infty} \frac{1}{2^r} \frac{1}{1 + \|f_r\|_{\infty}} \left| \sum_{\tilde{\alpha} \in \xi_1} \int f_r(x_1, \dots, x_{k_r}) \prod_{i=1}^{k_r} \alpha(dx_i) - \sum_{\tilde{\alpha} \in \xi_2} \int f_r(x_1, \dots, x_{k_r}) \prod_{i=1}^{k_r} \alpha(dx_i) \right|,$$

for a countable sequence of functions $\{f_r(x_1, x_2, \dots, x_{k_r})\}_{r \in \mathbb{N}}$ which is dense in \mathcal{F} . Here is our first main result.

THEOREM 3.1. \mathbf{D} is a metric on $\tilde{\mathcal{X}}$.

PROOF. Note that to prove \mathbf{D} is a metric the only nontrivial part that we need to show is that, two collections ξ_1 and ξ_2 are identical if $\mathbf{D}(\xi_1, \xi_2) = 0$. For this, it is enough to show if for every $k \geq 2$ and every f in \mathcal{F}_k ,

$$(18) \quad \sum_{\tilde{\alpha} \in \xi_1} \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) = \sum_{\tilde{\alpha} \in \xi_2} \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i),$$

then $\xi_1 = \xi_2$. We prove this into three steps.

Step 1. First, we show that, if (18) holds, then for every integer $r \geq 1$,

$$(19) \quad \sum_{\tilde{\alpha} \in \xi_1} \left\{ \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \right\}^r = \sum_{\tilde{\alpha} \in \xi_2} \left\{ \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \right\}^r.$$

This is certainly true for $r = 1$. For $r = 2$, we take a sequence g_N of functions of $2k$ variables defined by

$$g_N(x_1, x_2, \dots, x_{2k}) = f(x_1, x_2, \dots, x_k) f(x_{k+1}, x_{k+2}, \dots, x_{2k}) \varphi(N^{-1}(x_1 - x_{k+1})),$$

where $0 \leq \varphi \leq 1$ is equal to 1 inside a ball of radius 1 and is truncated smoothly to be 0 outside a ball of radius 2. Letting $N \rightarrow \infty$, for any $\alpha \in \mathcal{M}_{\leq 1}$, by the bounded convergence theorem,

$$\int g_N(x_1, \dots, x_{2k}) \prod_{i=1}^{2k} \alpha(dx_i) \rightarrow \left\{ \int f(x_1, x_2, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \right\}^2,$$

and we obtain

$$\sum_{\tilde{\alpha} \in \xi_1} \left\{ \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \right\}^2 = \sum_{\tilde{\alpha} \in \xi_2} \left\{ \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \right\}^2.$$

The general case for any $r \in \mathbb{N}$ follows from a similar argument.

Step 2. We note that if (19) holds for every $r \in \mathbb{N}$, we can identify for each $\alpha \in \mathcal{M}_{\leq 1}$ the values of

$$(20) \quad \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i),$$

for $f \in \mathcal{F}_k$ and $k \geq 2$. It follows that if (18) holds for any two elements ξ_1 and ξ_2 , then for every $f \in \mathcal{F}$, the list of values (20) for $\tilde{\alpha} \in \xi_1$ is the same as the list from ξ_2 .

However, this is not enough. We need to show that if (18) holds for any two elements ξ_1 and ξ_2 , then every $\tilde{\alpha} \in \xi_1$ occurs in ξ_2 with the same multiplicity (see Remark 3 for the definition of multiplicity of an orbit).

Let us denote by $S(f, \xi)$ the set of values of $\Lambda(f, \tilde{\mu})$ as $\tilde{\mu}$ varies over ξ . We have matched for $\xi = \xi_1$ and ξ_2 the set of values $S(f, \xi_1)$ and $S(f, \xi_2)$. The next step is to show that if we pick an orbit $\tilde{\mu}_1$ in ξ_1 , the set of values of $\Lambda(f, \tilde{\mu}_1)$ can actually be matched with the set of values $\Lambda(f, \tilde{\mu}_2)$ of some $\tilde{\mu}_2 \in \xi_2$ i.e. a single choice $\tilde{\mu}_2 \in \xi_2$ can be made to work for all $f \in \bigcup_k \mathcal{F}_k$. In other words, if we define for each $\tilde{\mu}_1 \in \xi_1$ and $\tilde{\mu}_2 \in \xi_2$

$$C_k(\tilde{\mu}_1, \tilde{\mu}_2) = \{f \in \mathcal{F}_k : \Lambda(f, \tilde{\mu}_1) = \Lambda(f, \tilde{\mu}_2)\},$$

then we have for each $\tilde{\mu}_1$

$$\mathcal{F}_k = \bigcup_{\tilde{\mu}_2 \in \xi_2} C_k(\tilde{\mu}_1, \tilde{\mu}_2).$$

Each $C_k \subset \mathcal{F}_k$ is a closed subset of a complete metric space and we have a countable union. By the Baire category theorem, at least one C_k has an interior. But if two linear functionals agree on an open set they agree everywhere. Therefore, there exists $\mu_2 \in \xi_2$ such that

$$(21) \quad \mathcal{F}_k = C_k(\tilde{\mu}_1, \mu_2).$$

The choice of μ_2 may still depend on k . We need to show that (21) holds for some μ_2 for all k . We note that any function $f(x_1, x_2, \dots, x_{k-1}) \in \mathcal{F}_{k-1}$ is a limit of

$$g(x_1, x_2, \dots, x_{k-1}, x_k) = f(x_1, x_2, \dots, x_{k-1})\varphi(x_{k-1} - x_k) \in \mathcal{F}_k$$

as the continuous function φ with compact support tends boundedly to 1. Therefore, if $\Lambda(f, \tilde{\mu}_1) = \Lambda(f, \mu_2)$ on \mathcal{F}_k , they agree on \mathcal{F}_{k-1} as well. In particular if $\Lambda(f, \tilde{\mu}_1) = \Lambda(f, \mu_2)$ on \mathcal{F}_k for infinitely many values of k , then they agree for all values of k . We note that by allowing φ to tend to 1, if $\Lambda(f, \tilde{\mu}_1) = \Lambda(f, \mu_2)$ on \mathcal{F}_2 , then μ_1 and μ_2 have the same mass. Assuming the mass to be positive, there can only be a finite number of possibilities for μ_2 since the total sum is at most 1. There is then a μ_2 that works for an infinite number of values of k and consequently for all k . We can then peel off matching pairs and proceed with what is left. If we are careful to remove at each stage measures with the largest masses from ξ_1 and ξ_2 , we will exhaust both ξ_1 and ξ_2 (it may take a countable number of steps).

Step 3. Now we have to recover the orbit of $\mu \in \mathcal{M}_{\leq 1}(\mathbb{R}^d)$ from the value

$$\int f(x_1, \dots, x_k) \prod_{i=1}^k \mu(dx_i)$$

for $f \in \mathcal{F}_k$. We can let f converge boundedly to $\exp\{\sum_{i=1}^k \sqrt{-1}\langle t_i, x_i \rangle\}$ provided $\sum_i t_i = 0$. In other words we can determine for the characteristic functions $\{\phi(t)e^{\sqrt{-1}\langle t, a \rangle}\}$ of $\tilde{\alpha} \in \mathcal{X}$, the value of $\prod_{i=1}^k \phi(t_i)$ for all $\{t_i\}$ with $\sum_i t_i = 0$.

The following calculation will complete the proof. Let $\phi(\cdot)$ and $\psi(\cdot)$ be two characteristic functions such that $\prod_{i=1}^k \phi(t_i) = \prod_{i=1}^k \psi(t_i)$ for all $\{t_i\}$ with $\sum_i t_i = 0$. In particular, $|\phi(t)|^2 = \phi(t)\phi(-t) = \psi(t)\psi(-t) = |\psi(t)|^2$. Let $G = \{t : |\phi(t)| = |\psi(t)| \neq 0\}$. Write $\phi(t) = \psi(t)\chi(t)$ on G . G is a symmetric open set containing 0. For any k and $t_1, \dots, t_k \in G$ such that $\sum_{i=1}^k t_i = \tau \in G$, we have

$$\prod_{i=1}^k \chi(t_i)\chi(-\tau) = 1.$$

Noting that $\chi(\tau) = \overline{\chi(-\tau)}$, we find that $\chi(t_1 + t_2 + \dots + t_k) = \prod_{i=1}^k \chi(t_i)$ provided, $\{t_i\}$ as well as $t_1 + \dots + t_k$ are all in G which contains a neighborhood of 0. It is now standard to show that for some $a \in \mathbb{R}^d$, $\chi(t) = e^{\sqrt{-1}\langle a, t \rangle}$ near 0 and since $\chi(kt) = (\chi(t))^k$ the proof is complete. \square

3.2. *Completion under the metric \mathbf{D} and the compactification.* Henceforth, the metric \mathbf{D} will define the topology on the space $\tilde{\mathcal{X}}$. Recall that the space of orbits $\tilde{\mathcal{M}}_1$ is canonically embedded in $\tilde{\mathcal{X}}$.

THEOREM 3.2. *The set of orbits $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$ is dense in $\tilde{\mathcal{X}}$. Furthermore, given any sequence $(\tilde{\mu}_n)_n$ in $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$, there is a subsequence that converges to a limit in $\tilde{\mathcal{X}}$. Hence, $\tilde{\mathcal{X}}$ is a compactification of $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$. It is then also the completion under the metric \mathbf{D} of the totally bounded space $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$.*

PROOF. We prove the theorem in two main steps.

Step 1. First, we show that $\tilde{\mathcal{M}}_1$ is dense in $\tilde{\mathcal{X}}$. Given any $\xi = \{\tilde{\alpha}_i : i \in I\} \in \tilde{\mathcal{X}}$, we would like to have a sequence $(\tilde{\mu}_n)_n$ in $\tilde{\mathcal{M}}_1$ which converges to $\xi \in \tilde{\mathcal{X}}$. This can be done if we take “distant shifts” of μ_n weighted by corresponding masses p_i of α_i . Any remaining mass $1 - \sum_i p_i$ can be filled by a Gaussian with a large variance (leading to “total disintegration” of mass $1 - \sum_i p_i$). The convex combination of all these measures will approximate ξ in $\tilde{\mathcal{X}}$.

Indeed, let $\xi = \{\tilde{\alpha}_i : i \in I\} \in \tilde{\mathcal{X}}$ be given. If it is an infinite collection, then for every $\varepsilon > 0$ we can pick a finite sub-collection $\{\alpha_1, \dots, \alpha_n\}$ such that the remaining

total masses $\sum_{j>n} \alpha_j(\mathbb{R}^d)$ add up to at most $\varepsilon > 0$. Since for any $\alpha \in \mathcal{M}_{\leq 1}$ and $f \in \mathcal{F}_k$,

$$\int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha(dx_i) \leq \|f\|_\infty (\alpha(\mathbb{R}^d))^k \leq \|f\|_\infty \alpha(\mathbb{R}^d),$$

we infer that

$$(22) \quad \sum_{j>n} \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha_j(dx_i) \leq \|f\|_\infty \sum_{j>n} \alpha_j(\mathbb{R}^d) \leq \varepsilon \|f\|_\infty.$$

Let us denote by $p_j = \alpha_j(\mathbb{R}^d)$ for $j = 1, \dots, n$ and choose spatial points $a_1, \dots, a_n \in \mathbb{R}^d$ so that $\inf_{i \neq j} |a_i - a_j| \rightarrow \infty$. Also, for any $M > 0$, let λ_M be a Gaussian in \mathbb{R}^d with mean $\mathbf{0} \in \mathbb{R}^d$ and covariance matrix $M\mathbf{Id} \in \mathbb{R}^{d \times d}$. Since the family of measures $\{\lambda_M\}_{M>0}$ totally disintegrates, by Lemma 2.3 and (13), for any $k \geq 2$ and $f \in \mathcal{F}_k$,

$$(23) \quad \lim_{M \rightarrow \infty} \int f(x_1, \dots, x_k) \prod_{i=1}^k \lambda_M(dx_i) = 0.$$

Then for the convex combination

$$(24) \quad \mu_n^{a_1, \dots, a_n, M} = \mu_n := \sum_{j=1}^n \alpha_j \star \delta_{a_j} + \left(1 - \sum_{j=1}^n p_j\right) \lambda_M,$$

we conclude that, for any $k \geq 2$ and $f \in \mathcal{F}_k$,

$$\int f(x_1, \dots, x_k) \prod_{i=1}^k \mu_n(dx_i) \longrightarrow \sum_{j=1}^n \int f(x_1, \dots, x_k) \prod_{i=1}^k \alpha_j(dx_i)$$

as $\inf_{i \neq j} |a_i - a_j| \rightarrow \infty$ and $M \uparrow \infty$, by (15) and (23) (masses that are far away from each other do not interact and masses that are too thinly spread do not count). Therefore, by (22) and the definition of the metric \mathbf{D} [recall (17)], the sequence of orbits $(\tilde{\mu}_n)_n$ converges to ξ in $\tilde{\mathcal{X}}$.

Step 2. We show that any sequence $(\tilde{\mu}_n)_n$ in $\tilde{\mathcal{X}}$ has a subsequence that converges to some $\xi \in \tilde{\mathcal{X}}$. We need to collect some facts.

Let $\mu \in \mathcal{M}_{\leq 1}(\mathbb{R}^d)$. The *concentration function* of μ is defined as,

$$(25) \quad q_\mu(r) = \sup_{x \in \mathbb{R}^d} \mu(B(x, r)),$$

for any $r > 0$. Then $\lim_{r \rightarrow \infty} q_\mu(r) = \mu(\mathbb{R}^d)$.

If $(\mu_n)_n$ is a sequence in $\mathcal{M}_{\leq 1}(\mathbb{R}^d)$, and $q_n(r)$ is the concentration function of μ_n , we can, by choosing a subsequence (which we suppress in the notation) if needed, assume that for any $r > 0$,

$$\lim_{n \rightarrow \infty} q_n(r) = q(r),$$

exists and also

$$\lim_{n \rightarrow \infty} \mu_n(\mathbb{R}^d) = p \in [0, 1].$$

If $q = \lim_{r \uparrow \infty} q(r)$, then always $q \leq p$.

If $q = 0$, we have for every $r > 0$

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \mu_n(B(x, r)) = \lim_{n \rightarrow \infty} q_n(r) = 0$$

and hence by Lemma 2.3, (13) and the definition of the metric \mathbf{D} [recall (17)], $\tilde{\mu}_n \rightarrow 0$ in $\tilde{\mathcal{X}}$.

If $q > 0$, then taking a suitable translation $a_n \in \mathbb{R}^d$, we can assume that $\lambda_n = \mu_n \star \delta_{a_n}$ satisfies, for some $r > 0$,

$$(26) \quad \lambda_n(B(0, r)) \geq q/2,$$

for all sufficiently large n . Let us assume, by choosing a subsequence if needed, $\lambda_n \rightharpoonup \alpha$. Then $\alpha(\mathbb{R}^d) \geq \frac{q}{2}$. According to Lemma 2.2, we can express $\lambda_n = \alpha_n + \beta_n$ where $\beta_n \rightharpoonup 0$ and $\alpha_n \rightharpoonup \alpha$. Lemma 2.4 implies that for $V \in \mathcal{F}_2$

$$\lim_{n \rightarrow \infty} \int V(x - y) \alpha_n(dx) \beta_n(dy) = 0.$$

This property is valid after translating back by δ_{-a_n} and μ_n has the same decomposition in terms of the shifted $\alpha_n \star \delta_{-a_n}$ and $\beta_n \star \delta_{-a_n}$. We will denote them again by α_n and β_n . We remark that if $q = p$, then $\lambda_n = \mu_n \star \delta_{a_n}$ converges weakly to α and β_n can be taken to be 0. To see this, choose $r > 0$ so that (26) holds. Furthermore, note that given any $\varepsilon > 0$, there are translations $b_{n,\varepsilon}$ such that, for some r_ε , $(\mu_n \star \delta_{b_n})[B(0, r_\varepsilon)] \geq p - \varepsilon$ for large enough n . The sets $B(-a_n, r)$ and $B(-b_n, r_\varepsilon)$ can not be disjoint, because if they were, there combined total mass would exceed p [recall (26)]. Therefore, $|a_n - b_n| \leq r + r_\varepsilon$. This implies $B(-a_n, r + 2r_\varepsilon) \supset B(-b_n, r_\varepsilon)$ and $\lambda_n[B(0, r + 2r_\varepsilon)] \geq p - \varepsilon$. This shows that λ_n is a tight family of measures and (choosing a subsequence if needed) $\lambda_n \rightharpoonup \alpha$ for some $\alpha \in \mathcal{M}_{\leq 1}(\mathbb{R}^d)$ and β_n can be taken as 0. Hence, again by definition of the metric \mathbf{D} [recall (17)], $\tilde{\mu}_n \rightarrow \tilde{\alpha}$ in $\tilde{\mathcal{X}}$.

Let us now start with a sequence $(\mu_n)_n$ in $\mathcal{M}_1(\mathbb{R}^d)$. We want to prove that the sequence $(\tilde{\mu}_n)_n$ in $\tilde{\mathcal{X}}$ has a subsequence that converges to some $\xi \in \tilde{\mathcal{X}}$. Hence, to begin with $p = 1$ and $0 \leq q \leq 1$. By the remarks made above, if $q = 0$, then $\tilde{\mu}_n \rightarrow 0$ in $\tilde{\mathcal{X}}$. If $q = 1$, then $\alpha_n = \mu_n$ and $\tilde{\mu}_n \rightarrow \tilde{\alpha}$ in $\tilde{\mathcal{X}}$.

If $0 < q < 1$, we can, for some sequence $(a_n)_n \subset \mathbb{R}^d$, represent $\mu_n = \alpha_n + \beta_n$. α_n so that:

-

$$\alpha_n \star \delta_{a_n} \rightharpoonup \alpha.$$

- For every $V \in \mathcal{F}_2$,

$$\lim_{n \rightarrow \infty} \int V(x_1 - x_2) \alpha_n(dx_1) \beta_n(dx_2) = 0.$$

- For every $r > 0$,

$$\lim_{n \rightarrow \infty} q_{\beta_n}(r) \leq \min \left\{ 1 - \frac{q}{2}, q \right\}.$$

The last inequality requires a remark. Since $\beta_n \leq \mu_n$ we have $q_{\beta_n}(r) \leq q_{\mu_n}(r)$ for every r . Mass of $\frac{q}{2}$ has been removed in the limit from μ_n by (26). What is left can in the limit have mass at most $1 - \frac{q}{2}$.

We repeat the procedure with β_n . Either the process goes on forever or terminates at some finite stage. If it terminates at a finite stage, we would have the decomposition

$$(27) \quad \mu_n = \sum_{j=1}^k \alpha_n^{(j)} + \gamma_n \quad k \in \mathbb{N},$$

that will satisfy:

- For $j = 1, \dots, k$,

$$\lim_{n \rightarrow \infty} \alpha_n^{(j)} * a_n^{(j)} \Rightarrow \alpha_j.$$

- For $i \neq j$ and $V \in \mathcal{F}_2$,

$$\lim_{n \rightarrow \infty} \int V(x_1 - x_2) \alpha_n^{(i)}(dx_1) \alpha_n^{(j)}(dx_2) = 0.$$

- For every $r > 0$, $q_{\gamma_n}(r) \rightarrow 0$ and

$$\lim_{n \rightarrow \infty} \int V(x_1 - x_2) \gamma_n(dx_1) \gamma_n(dx_2) = 0.$$

Clearly $\tilde{\mu}_n$ converges to $\xi = \{\tilde{\alpha}_1, \dots, \tilde{\alpha}_k\}$ in $\tilde{\mathcal{X}}$.

If the process continues forever, we have for each $k \in \mathbb{N}$ a decomposition as above. Inductively, starting from $\beta_{n,0} = \mu_n$, we define according to Lemma 2.2 $\beta_{n,j} = \alpha_{n,j+1} + \beta_{n,j+1}$ so that $\alpha_{n,j} \Rightarrow \alpha_j$. Let $p_j = \lim_{n \rightarrow \infty} \beta_j(\mathbb{R}^d)$ and $q_j = \lim_{r \rightarrow \infty} \lim_{n \rightarrow \infty} q_{\beta_{n,j}}(r)$. Since $\alpha_j(\mathbb{R}^d) \geq \frac{q_j}{2}$, and $\sum_j \alpha_j(\mathbb{R}^d) \leq 1$, it follows that $q_j \rightarrow 0$ as $j \rightarrow \infty$. Fix any $F \in \mathcal{F}_k$. Then, proceeding inductively in j ,

$$\begin{aligned} \int F(x_1, \dots, x_k) \mu_n(dx_1) \cdots \mu_n(dx_k) &= \sum_{i=1}^j \int F(x_1, \dots, x_k) \alpha_{n,i}(dx_1) \cdots \alpha_{n,i}(dx_k) \\ &\quad + \int F(x_1, \dots, x_k) \beta_{n,j}(dx_1) \cdots \beta_{n,j}(dx_k). \end{aligned}$$

Since $q_j \rightarrow 0$ and the orbits $\tilde{\alpha}_{n,j}$ converge to $\tilde{\alpha}_j$ in $\tilde{\mathcal{X}}$, the theorem is proved. \square

We end this section with an immediate corollary which will be of use later.

COROLLARY 3.3. *Let $(\tilde{\mu}_n)_n$ be a sequence in $\tilde{\mathcal{X}}$ so that $\tilde{\mu}_n \rightarrow \xi = \{\tilde{\alpha}_j\} \in \tilde{\mathcal{X}}$. Then, for any $V \in \mathcal{F}_2$,*

$$\lim_{n \rightarrow \infty} \int \int_{\mathbb{R}^d \times \mathbb{R}^d} V(x - y) \mu_n(dx) \mu_n(dy) = \sum_j \int \int_{\mathbb{R}^d \times \mathbb{R}^d} V(x - y) \alpha_j(dx) \alpha_j(dy).$$

In other words, the functional

$$H(\tilde{\mu}) = \int \int_{\mathbb{R}^d \times \mathbb{R}^d} V(x - y) \mu(dx) \mu(dy) \quad \mu \in \mathcal{M}_1(\mathbb{R}^d),$$

is continuous on $\tilde{\mathcal{X}}$.

4. Large deviation principles in the compact space $\tilde{\mathcal{X}}$. Recall that we started with Wiener measure \mathbb{P} on $\Omega = C[[0, \infty); \mathbb{R}^d]$ corresponding to the d -dimensional Brownian motion W starting from the origin with

$$L_t(A) = \frac{1}{t} \int_0^t \mathbb{1}_A(W(s)) ds \quad A \subset \mathbb{R}^d$$

denoting its normalized occupation measure until time t . Note that L_t maps

$$(28) \quad \Omega \rightarrow \mathcal{M}_1(\mathbb{R}^d)$$

inducing a probability distribution on $\mathcal{M}_1(\mathbb{R}^d)$. Classical large deviation principle [2] states that the family of these distributions satisfies a “weak” large deviation principle in the space probability measures on $\mathcal{M}_1(\mathbb{R}^d)$ equipped with the weak topology with a rate function I . More precisely, for every compact subset $K \subset \mathcal{M}_1(\mathbb{R}^d)$,

$$(29) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}(L_t \in K) \leq - \inf_{\mu \in K} I(\mu)$$

and for every open subset $G \subset \mathcal{M}_1(\mathbb{R}^d)$

$$(30) \quad \liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{P}(L_t \in G) \geq - \inf_{\mu \in G} I(\mu),$$

where I is the rate function given by

$$(31) \quad I(\mu) = \begin{cases} \frac{1}{2} \|\nabla f\|_2^2, & \text{if } f = \sqrt{\frac{d\mu}{dx}} \in H^1(\mathbb{R}^d), \\ \infty, & \text{else.} \end{cases}$$

Here $H^1(\mathbb{R}^d)$ is the usual Sobolev space of square integrable functions with square integrable derivatives. Note that the function $\mu \mapsto I(\mu)$ is translation invariant and depends only on the orbit $\tilde{\mu}$. Furthermore, this map is convex and homogenous of degree 1.

We say that a family of measures satisfies a “strong” large deviation principle, or simply a *large deviation principle* (LDP) if the upper bound (29) holds for all closed sets.

Note that we also have an extension of (28) via

$$\Omega \rightarrow \mathcal{M}_1(\mathbb{R}^d) \rightarrow \tilde{\mathcal{M}}_1(\mathbb{R}^d) \subset \tilde{\mathcal{X}}$$

which induces a probability distribution \mathbb{Q}_t of \tilde{L}_t on $\tilde{\mathcal{X}}$. Our second main result gives a large deviation principle for \mathbb{Q}_t on $\tilde{\mathcal{X}}$ with the rate function

$$(32) \quad \tilde{I}(\xi) = \sum_{\tilde{\alpha} \in \xi} I(\tilde{\alpha}),$$

where

$$I(\tilde{\alpha}) = I(\alpha),$$

where I is defined in (31) and α is any arbitrary element of the orbit $\tilde{\alpha}$ (recall that I is translation invariant). We remark that although I is defined in (31) only on probability measures $\mathcal{M}_1(\mathbb{R}^d)$, the definition canonically extends to sub-probability measures $\mathcal{M}_{\leq 1}(\mathbb{R}^d)$. Here is our second main result.

THEOREM 4.1. *The family of measures $\{\mathbb{Q}_t\}_t$ on the compact metric space $\tilde{\mathcal{X}}$ equipped with the metric \mathbf{D} satisfies a large deviation principle with the rate function $\tilde{I}(\xi)$ defined in (32).*

We split the proof into three main steps. First, we prove that the function \tilde{I} is lower semicontinuous on $\tilde{\mathcal{X}}$.

LEMMA 4.2 (Lower semicontinuity). *If $\xi_n \rightarrow \xi$ in $\tilde{\mathcal{X}}$, then*

$$\liminf_{n \rightarrow \infty} \tilde{I}(\xi_n) \geq \tilde{I}(\xi).$$

PROOF. Let us first consider the case where, for each $n \in \mathbb{N}$, ξ_n consists of a single orbit $\tilde{\mu}_n$ and the limit ξ is a finite or countable collection $\{\tilde{\alpha}_i\}$ arranged so that their masses $\{p_i\}$ form a non-increasing sequence. Given $\varepsilon > 0$, it is then possible to write [recall (27)]

$$\mu_n = \sum_{i=1}^k \alpha_n^{(i)} + \beta_n$$

for some $k \in \mathbb{N}$ such that the following properties hold: For each $i = 1, \dots, k$, there are sequences $\{a_n^{(i)}\}_n \subset \mathbb{R}^d$ such that

$$\alpha_n^{(i)} * \delta_{a_n^{(i)}} \Rightarrow \alpha_i \in \tilde{\alpha}_i,$$

$$\lim_{n \rightarrow \infty} \inf_{i \neq j} |a_n^{(i)} - a_n^{(j)}| = \infty,$$

$$\lim_{n \rightarrow \infty} \int V(x - y) \alpha_n^{(i)}(dx) \beta_n(dy) = 0,$$

$$\limsup_{n \rightarrow \infty} \int V(x - y) \beta_n(dx) \beta_n(dy) \leq 2\varepsilon,$$

for all $V \in \mathcal{F}_2$. In particular, since for each $i = 1, \dots, k$, $\alpha_n^{(i)}$ is weakly convergent, they are a tight sequence and therefore $\alpha_n^{(i)}$ is concentrated near $-a_n^{(i)}$. We can find a smooth cut-off function $\varphi(x)$ which is 1 in the unit ball, 0 outside a ball of radius 2 and smoothly varies in between. In particular, $0 \leq \varphi \leq 1$. For $r_n > 0$ to be suitably chosen later, we will have a partition of unity by setting

$$1 = \sum_{i=1}^k \left\{ \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right\}^2 + \left[1 - \sum_{i=1}^k \left\{ \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right\}^2 \right] \quad r_n > 0.$$

We can assume that $I(\mu_n) < \infty$ for each $n \in \mathbb{N}$ (since otherwise there is nothing to prove) and hence $\mu_n(dx) = f_n(x) dx$ and $f_n \in H^1(\mathbb{R}^d)$. If $g_n = \sqrt{f_n}$ and $\frac{1}{2} \int_{\mathbb{R}^d} |\nabla g_n|^2 dx \leq \ell$, we need to prove that $\alpha_1, \dots, \alpha_k$ are all absolutely continuous with densities f^1, \dots, f^k and

$$\sum_{i=1}^k I(f^{(i)}) \leq \ell.$$

We define, for any $i = 1, \dots, k$,

$$\begin{aligned} f_n^{(i)}(x) &= f_n(x) \left\{ \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right\}^2 \\ &= \left\{ g_n(x) \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right\}^2 \end{aligned}$$

and we let $r_n \rightarrow \infty$ in such a way that $2r_n \leq \min_{i \neq j} |a_n^{(i)} - a_n^{(j)}|$. Then $f_n^{(i)}(x) dx \Rightarrow \alpha_i$ for $i = 1, 2, \dots, k$ and

$$I(f_n^{(i)}) = \frac{1}{2} \int \left| \nabla g_n(x) \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) + \frac{1}{r_n} g_n(x) (\nabla \varphi)\left(\frac{x + a_n^{(i)}}{r_n}\right) \right|^2 dx.$$

Since $r_n \rightarrow \infty$, φ and $\nabla \varphi$ are uniformly bounded and the integrals $\int |g_n(x)|^2 dx$ and $\int |\nabla g_n|^2 dx$ are bounded, only the first term in the integral counts. Since the functions

$$\left\{ \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right\}_{i=1, \dots, k}$$

do not overlap and $0 \leq \varphi \leq 1$, we infer

$$\begin{aligned} \sum_{i=1}^k \frac{1}{2} \int \left| \nabla g_n(x) \varphi\left(\frac{x + a_n^{(i)}}{r_n}\right) \right|^2 dx &\leq \frac{1}{2} \int |\nabla g_n(x)|^2 dx \\ &= \frac{1}{2} \int |\nabla \sqrt{f_n(x)}|^2 dx = I(f_n) \leq \ell. \end{aligned}$$

This implies that any weak limit α_i of $f_n^{(i)} dx$ has a density $f^{(i)}$ and $\sum_{i=1}^k I(f^{(i)}) \leq \ell$.

Finally, if ξ_n consists of multiple orbits $\{\xi_n^{(i)}\}_i$ with $\sum_i \tilde{I}(\xi_n^{(i)}) \leq \ell$, we can choose subsequences such that, for each i , $\xi_n^{(i)}$ has a limit which is a collection $\xi^{(i)}$ of orbits $\{\tilde{\alpha}_j^{(i)}\}_j$. The last step implies, for each i , $\sum_j \tilde{I}(\tilde{\alpha}_j^{(i)}) \leq \ell^{(i)}$ where $\ell^{(i)} = \liminf_{n \rightarrow \infty} \tilde{I}(\xi_n^{(i)})$. Hence,

$$I(\xi) = \sum_i \ell^{(i)} \leq \liminf_{n \rightarrow \infty} \sum_i \tilde{I}(\xi_n^{(i)}) \leq \ell.$$

This proves the lemma. \square

Next, we derive the large deviation lower bound for \mathbb{Q}_t on $\tilde{\mathcal{X}}$. This is easily done given the translation invariance, convexity and homogeneity of I and the denseness of the space $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$ in $\tilde{\mathcal{X}}$.

LEMMA 4.3 (Lower bound). *For any open set G in $\tilde{\mathcal{X}}$,*

$$(33) \quad \liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{Q}_t(G) \geq - \inf_{\xi \in G} \tilde{I}(\xi).$$

PROOF. For (33) it is enough to prove, given $\xi \in \tilde{\mathcal{X}}$ with $\tilde{I}(\xi) < \infty$,

$$(34) \quad \liminf_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{Q}_t(U) \geq -\tilde{I}(\xi),$$

for any neighborhood $U \ni \xi$.

We claim that any $\xi \in \tilde{\mathcal{X}}$ with $\tilde{I}(\xi) < \infty$ can be approximated by $\xi_n \in \tilde{\mathcal{X}}$ such that

$$(35) \quad \limsup_{n \rightarrow \infty} \tilde{I}(\xi_n) \leq \tilde{I}(\xi).$$

Indeed, recall from step-1 of the proof of Theorem 3.2 that $\tilde{\mathcal{M}}_1$ is dense in $\tilde{\mathcal{X}}$ and $\xi = \{\tilde{\alpha}_j\} \in \tilde{\mathcal{X}}$ can be approximated by the sequence $(\tilde{\mu}_n)_n$ in $\tilde{\mathcal{M}}_1$, where, as constructed in (24),

$$\mu_n := \sum_{j=1}^n \alpha_j \star \delta_{a_j} + \left(1 - \sum_{j=1}^n p_j\right) \lambda_M \in \mathcal{M}_1(\mathbb{R}^d),$$

and λ_M is a Gaussian with mean vector $\mathbf{0}$ and covariance matrix $M\mathbf{Id}$. Furthermore, since $I(\cdot)$ on $\mathcal{M}_1(\mathbb{R}^d)$ is translation invariant, homogeneous of degree 1 and convex, it is also sub-additive on $\mathcal{M}_{\leq 1}(\mathbb{R}^d)$. Then,

$$\begin{aligned} I(\mu_n) &\leq \sum_{j=1}^n I(\alpha_j \star \delta_{a_j}) + \left(1 - \sum_{j=1}^n p_j\right) I(\lambda_M) \\ &= \sum_{j=1}^n I(\alpha_j) + \left(1 - \sum_{j=1}^n p_j\right) I(\lambda_M) \end{aligned}$$

$$\begin{aligned} &\leq \tilde{I}(\xi) + I(\lambda_M) \\ &= \tilde{I}(\xi) + \frac{1}{M}. \end{aligned}$$

Since we can choose M to depend on n , make it arbitrarily large and take (ξ_n) to be the single orbit sequence $(\tilde{\mu}_n)$, (35) is proved. The desired lower bound (34) now follows from the large deviation lower bound (30) of the distribution of L_t on $\mathcal{M}_1(\mathbb{R}^d)$. \square

Finally, we turn to the large deviation upper bound for \mathbb{Q}_t .

PROPOSITION 4.4 (Upper bound of Theorem 4.1). *For any closed set F in $\tilde{\mathcal{X}}$,*

$$(36) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{Q}_t(F) \leq - \inf_{\xi \in F} \tilde{I}(\xi).$$

Let \mathcal{U} be the space of functions of the form $u = c + v$ where v is a smooth nonnegative function with compact support on \mathbb{R}^d and $c > 0$ is a positive constant. Let $\varphi(x)$ be a smooth function satisfying $0 \leq \varphi(x) \leq 1$, $\varphi(x) = 1$ inside the unit ball and $\varphi(x) = 0$ outside the ball of radius 2. For any $k \geq 1$, $R > 0$, $u_1, \dots, u_k \in \mathcal{U}$ and $a_1, \dots, a_k \in \mathbb{R}^d$ and $c > 0$ consider the function

$$(37) \quad g(x) = g(k, R, c, a_1, \dots, a_k, x) = c + \sum_{i=1}^k u_i(x + a_i) \varphi\left(\frac{x + a_i}{R}\right)$$

and define $F : \Omega \rightarrow \mathbb{R}$ by setting

$$\begin{aligned} (38) \quad F(u_1, \dots, u_k, c, R, t, \omega) &= \sup_{\substack{a_1, \dots, a_k \\ \inf_{i \neq j} |a_i - a_j| \geq 4R}} \frac{1}{t} \int_0^t \frac{-1/2 \Delta g(W(s))}{g(W(s))} ds \\ &= \sup_{\substack{a_1, \dots, a_k \\ \inf_{i \neq j} |a_i - a_j| \geq 4R}} \int_{\mathbb{R}^d} \frac{-1/2 \Delta g(x)}{g(x)} L_t(dx). \end{aligned}$$

Since the last expression depends only on the image \tilde{L}_t of L_t in $\tilde{\mathcal{X}}$, we write

$$\begin{aligned} (39) \quad \tilde{F}(u_1, \dots, u_k, c, R, \tilde{L}_t) &= \sup_{\substack{a_1, \dots, a_k \\ \inf_{i \neq j} |a_i - a_j| \geq 4R}} \int_{\mathbb{R}^d} \frac{-1/2 \Delta g(x)}{g(x)} L_t(dx) \\ &= F(u_1, \dots, u_k, c, R, t, \omega). \end{aligned}$$

We will need the next three lemmas to prove the upper bound. First, we prove that $\tilde{F}(\cdot)$ grows only sub-exponentially as $t \rightarrow \infty$.

LEMMA 4.5. For any $k \geq 1, R > 0, u_1, \dots, u_k \in \mathcal{U}$ and $c > 0$,

$$\begin{aligned}
 & \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} \left\{ \exp \left\{ t \tilde{F}(u_1, \dots, u_k, c, R, \tilde{L}_t) \right\} \right\} \\
 (40) \quad & = \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} \left\{ \exp \left\{ t F(u_1, \dots, u_k, c, R, t, \omega) \right\} \right\} \\
 & \leq 0.
 \end{aligned}$$

PROOF. If it were not for the supremum over a_1, \dots, a_k this would be a simple consequence of Feynman–Kac formula. In fact, we first show that

$$(41) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} \left\{ \exp \left\{ \int_0^t -\frac{\Delta g(W(s))}{2g(W(s))} ds \right\} \right\} = 0.$$

Indeed, by the Feynman–Kac formula, the function

$$\Psi(t, x) = \mathbb{E}_x \left\{ g(W_t) \exp \left\{ \int_0^t -\frac{\Delta g(W_s)}{2g(W_s)} ds \right\} \right\}$$

satisfies the initial value problem

$$\begin{cases} \frac{\partial}{\partial t} \Psi = \frac{1}{2} \Delta \Psi(t, x) - \frac{\Delta g(x)}{2g(x)} \Psi(t, x), \\ \Psi(0, x) = g(x). \end{cases}$$

However, we clearly see that

$$\Psi(t, x) = g(x)$$

solves the above heat equation. Furthermore by definition [recall (37)],

$$g(x) \geq c.$$

Hence, we conclude

$$\begin{aligned}
 g(x) &= \mathbb{E}_x \left\{ g(W_t) \exp \left\{ \int_0^t -\frac{\Delta g(W_s)}{2g(W_s)} ds \right\} \right\} \\
 &\geq c \mathbb{E}_x \left\{ \exp \left\{ \int_0^t -\frac{\Delta g(W_s)}{2g(W_s)} ds \right\} \right\}
 \end{aligned}$$

and therefore,

$$(42) \quad \mathbb{E}_x \left\{ \exp \left\{ \int_0^t -\frac{\Delta g(W_s)}{2g(W_s)} ds \right\} \right\} \leq \frac{g(x)}{c}.$$

This proves (41). To handle the supremum over (a_1, \dots, a_k) inside the expectation we have to do a ‘‘course graining’’ argument.

First, we note that if the range of the Brownian motion in the time interval $[0, t]$ is r_t , once any $|a_i|$ exceeds $r_t + R$ it will no longer affect the value of g [again recall

the definition (37)]. We can therefore limit each a_i to the ball of radius $r_t + R$. But $P[r_t + R \geq t^2] \leq \exp[-c_1 t^3]$ and can be ignored. In other words, we can limit each a_i to the ball of radius t^2 .

Furthermore, the function $\frac{-1/2\Delta g(x)}{g(x)}$ is a uniformly continuous function of a_1, \dots, a_k and given any $\varepsilon > 0$, there is a $\delta > 0$ such that its oscillation in a box of size δ is at most ε . A ball of radius t^2 can be covered by $(\frac{t^2}{\delta})^{dk}$ such boxes. There is a set $K \subset (\mathbb{R}^d)^k$ of representatives of such boxes, a set of cardinality at most $(\frac{t^2}{\delta})^{dk}$ satisfying $|a_i - a_j| \geq 4R$ for all $i \neq j$.

Using the above two remarks, we can now estimate

$$\begin{aligned} & \mathbb{E} \left\{ \exp \left\{ \sup_{\substack{a_1, \dots, a_k \\ |a_i - a_j| \geq 4R \forall i \neq j}} \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} \\ & \leq \mathbb{E} \left\{ \exp \left\{ \sup_{\substack{|a_1| \leq t^2, \dots, |a_k| \leq t^2 \\ |a_i - a_j| \geq 4R \forall i \neq j}} \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} + e^{c_2 t} P \left(\sup_{0 \leq s \leq t} |W_s| \geq t^2 \right) \\ & \leq \mathbb{E} \left\{ \exp \left\{ \varepsilon t + \sup_{(a_1, \dots, a_k) \in K} \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} + \exp(c_2 t - c_1 t^3) \\ & \leq \mathbb{E} \left\{ \sum_{(a_1, \dots, a_k) \in K} \exp \left\{ \varepsilon t + \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} + \exp(c_2 t - c_1 t^3) \\ & \leq \left(\frac{t^2}{\delta} \right)^{dk} \sup_{(a_1, \dots, a_k) \in K} E \left\{ \exp \left\{ \varepsilon t + \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} \\ & \quad + \exp(c_2 t - c_1 t^3). \end{aligned}$$

Taking logarithm, dividing by t , passing to $\lim_{t \rightarrow \infty}$ and invoking (41), we obtain

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log E \left\{ \exp \left\{ \sup_{\substack{a_1, \dots, a_k \\ |a_i - a_j| \geq 4R \forall i \neq j}} \int_0^t \frac{-1/2\Delta g(W_s)}{g(W_s)} ds \right\} \right\} \leq \varepsilon,$$

and $\varepsilon > 0$ is arbitrary. (40) is proved. \square

LEMMA 4.6. *Let $(\tilde{\mu}_n)_n$ be sequence in $\tilde{\mathcal{X}}$ which converges to $\xi = \{\tilde{\alpha}_j\} \in \tilde{\mathcal{X}}$. For any $k \in \mathbb{N}$, $i = 1, \dots, k$ and $u_{i,\ell}(x) = u_i(x)\varphi(\frac{x}{R})$, where $u_i \in \mathcal{U}$, we have*

$$\begin{aligned} (43) \quad \liminf_{n \rightarrow \infty} \tilde{F}(u_1, \dots, u_k, c, R, \tilde{\mu}_n) & \geq \sum_{i=1}^k \int \frac{-(1/2\Delta u_{i,R})(x)}{c + u_{i,R}(x)} \alpha_i(dx) \\ & = \tilde{\Lambda}(\xi, R, c, u_1, \dots, u_k). \end{aligned}$$

PROOF. If $\tilde{\mu}_n \rightarrow \xi = \{\tilde{\alpha}_j\}$, then for $j = 1, 2, \dots, k$ we can again decompose μ_n as [recall (27)] $\mu_n = \sum_{j=1}^k \alpha_{n,j} + \beta_n$ with $\alpha_{n,j} * \delta_{a_{n,j}} \Rightarrow \alpha_j$ for a suitable choice of $a_{n,j}$ that satisfy $\lim_{n \rightarrow \infty} |a_{n,i} - a_{n,j}| = \infty$ for $i \neq j$. If n is large enough $|a_{n,i} - a_{n,j}| \geq 4R$ and the supports of $\{u_{i,R}\}_i$ are mutually disjoint. In particular with the choice of $a_i = -a_{n,i}$

$$\frac{-1/2\Delta g(k, R, c, a_1, \dots, a_k, x)}{g(k, R, c, a_1, \dots, a_k, x)} = \sum_{i=1}^k \frac{-(1/2\Delta u_{i,R})(x - a_{n,i})}{c + u_{i,R}(x - a_{n,i})}.$$

Because $\alpha_{n,j}$ is gets widely separated from β_n as well as $\alpha_{n,i}$ for $i \neq j$, it is clear that

$$\lim_{n \rightarrow \infty} \int \frac{-(1/2\Delta u_{i,R})(x - a_{n,i})}{c + u_{i,R}(x - a_{n,i})} \mu_n(dx) = \int \frac{-(1/2\Delta u_{i,R})(x)}{c + u_{i,R}(x)} \alpha_i(dx),$$

and the lemma follows. \square

LEMMA 4.7. With $\tilde{\Lambda}$ defined in (43) and \tilde{I} defined in (32), we have the identification

$$\tilde{I}(\xi) = \sup_{\substack{R, c > 0, k \in \mathbb{N}, \\ u_1, \dots, u_k \in \mathcal{U}}} \tilde{\Lambda}(\xi, R, c, u_1, \dots, u_k).$$

PROOF. Recall the definition of the classical rate function I from (31). For any $\alpha \in \mathcal{M}_{\leq 1}(\mathbb{R}^d)$, we can also identify I as

$$I(\alpha) = \sup_{\substack{u \in \mathcal{U} \\ c > 0}} \int \frac{-1/2\Delta u(x)}{c + u(x)} \alpha(dx).$$

Therefore for every $k \in \mathbb{N}$,

$$\sup_{\substack{c > 0, R > 0 \\ u_1, \dots, u_k \in \mathcal{U}}} \tilde{\Lambda}(\xi, c, R, u_1, \dots, u_k) = \sum_{i=1}^k I(\alpha_j)$$

and

$$\sup_{k \in \mathbb{N}} \sum_{j=1}^k I(\alpha_j) = \sum_{j=1}^{\infty} I(\alpha_j) = \tilde{I}(\xi). \quad \square$$

Now we come to the proof of the large deviation upper bound for \mathbb{Q}_t in $\tilde{\mathcal{X}}$.

PROOF OF PROPOSITION 4.4. $\tilde{\mathcal{X}}$ being compact, for (36), it is enough to prove (by the usual machinery of covering a compact space by finitely many balls

and invoking the union of events bound) that if $\xi \in \tilde{\mathcal{X}}$ and N_δ is a ball (as usual, in the metric \mathbf{D}) of radius δ around ξ , then

$$(44) \quad \limsup_{\delta \rightarrow 0} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{Q}_t(N_\delta) \leq -\tilde{I}(\xi).$$

Let \mathcal{H} be the space of maps $H : \tilde{\mathcal{M}}_1(\mathbb{R}^d) \rightarrow \mathbb{R}$ with the following properties: For each H there is a corresponding function $\Lambda_H : \tilde{\mathcal{X}} \rightarrow \mathbb{R}$ such that

$$(45) \quad \liminf_{\substack{\tilde{\mu} \in \tilde{\mathcal{M}}_1(\mathbb{R}^d) \\ \tilde{\mu} \rightarrow \xi \in \tilde{\mathcal{X}}}} H(\tilde{\mu}) \geq \Lambda_H(\xi)$$

and

$$(46) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}^{\mathbb{Q}_t} \{ \exp\{tH(\cdot)\} \} \leq 0.$$

Then again the properties of the decomposition (27), a routine application of Tchebycheff’s inequality, (45) and (46) show that, for any $H \in \mathcal{H}$,

$$\limsup_{\delta \rightarrow 0} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{Q}_t[N_\delta] \leq -\Lambda_H(\xi).$$

It is therefore enough to identify $\tilde{I}(\xi)$ as

$$(47) \quad \tilde{I}(\xi) = \sup_{H \in \mathcal{H}} \Lambda_H(\xi).$$

Recall the definition of \tilde{F} from (39). Then, for \mathcal{H} , by Lemma 4.5, Lemma 4.6 and Lemma 4.7, we can take the collection $\{\tilde{F}(u_1, \dots, u_k, c, R, \tilde{\mu})\}$ with $k \in \mathbb{N}$, $R, c > 0$, $u_1, \dots, u_k \in \mathcal{U}$ and $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and set $\Lambda_{\tilde{F}} = \tilde{\Lambda}$, with $\tilde{\Lambda}$ defined in (43). This proves (47) and hence Proposition 4.4. \square

5. Application: Localization of path measures with Coulomb interaction.

In this section, we come back to the problem we introduced in Section 1. Again, we consider the Wiener measure \mathbb{P} on $\Omega = C_0([0, \infty); \mathbb{R}^3)$ corresponding to a three dimensional Brownian motion $W = (W_t)_{t \geq 0}$ starting at the origin. Consider the transformed measure

$$d\hat{\mathbb{P}}_t = \frac{1}{Z_t} \exp\left\{ \frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} ds d\sigma \right\} d\mathbb{P},$$

where

$$Z_t = \mathbb{E} \left[\exp\left\{ \frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} ds d\sigma \right\} \right]$$

is the normalizing constant or the *partition function*. As mentioned before (see [4]),

$$(48) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \log Z_t = \sup_{\substack{\psi \in H^1(\mathbb{R}^3) \\ \|\psi\|_2=1}} \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} dx dy \frac{\psi^2(x)\psi^2(y)}{|x - y|} - \frac{1}{2} \|\nabla \psi\|_2^2 \right\},$$

and according to the classical result of Lieb (see [5]), this variational problem admits a maximizer ψ_0 which is radially symmetric and is unique up to translations. Let $d\mu_0 = \psi_0^2(x) dx$ define its probability distribution with $\tilde{\mu}_0$ the corresponding orbit in $\tilde{\mathcal{M}}_1(\mathbb{R}^3)$. We study the distribution

$$\hat{\mathbb{Q}}_t = \hat{\mathbb{P}}_t \tilde{L}_t^{-1}$$

on $\tilde{\mathcal{M}}_1(\mathbb{R}^3)$ of the orbit \tilde{L}_t of the normalized occupation measures L_t of the trajectory $\{W_s : 0 \leq s \leq t\}$ under the transformed measure $\hat{\mathbb{P}}_t$. Here is our next main result.

THEOREM 5.1 (The tube property under Coulomb interaction). *As probability measures on $\tilde{\mathcal{M}}_1(\mathbb{R}^3)$,*

$$\lim_{t \rightarrow \infty} \hat{\mathbb{Q}}_t = \delta_{\tilde{\mu}_0}$$

under the weak topology.

REMARK 4. Note that the topology on $\tilde{\mathcal{M}}_1(\mathbb{R}^3)$ is the same as weak convergence. As we shall see, the compactification $\tilde{\mathcal{X}}$ of $\tilde{\mathcal{M}}_1(\mathbb{R}^3)$ plays a role only in the proof of this theorem and not in its statement.

The proof involves the standard large deviation route. The function $\frac{1}{|x|}$ is unbounded and needs to be truncated to fit within the standard large deviation theory. We write

$$(49) \quad \frac{1}{|x|} = V_\varepsilon(x) + Y_\varepsilon(x)$$

with $V_\varepsilon(x) = (\varepsilon^2 + |x|^2)^{-1/2}$. The difference is given by

$$\begin{aligned} Y_\varepsilon(x) &= \frac{1}{|x|} - \frac{1}{\sqrt{\varepsilon^2 + |x|^2}} = \frac{\sqrt{\varepsilon^2 + |x|^2} - |x|}{|x|\sqrt{\varepsilon^2 + |x|^2}} \\ &= \frac{\varepsilon^2}{|x| + \sqrt{\varepsilon^2 + |x|^2}} \frac{1}{\sqrt{\varepsilon^2 + |x|^2}} \frac{1}{|x|} \\ &= \varepsilon^{-1} \varphi\left(\frac{x}{\varepsilon}\right), \end{aligned}$$

with

$$\phi(x) = \frac{1}{|x|} \frac{1}{\sqrt{1 + |x|^2}} \frac{1}{(|x| + \sqrt{1 + |x|^2})}.$$

We need the following lemma to control the difference.

LEMMA 5.2. For any $\lambda > 0$,

$$(50) \quad \limsup_{\varepsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E} \left[\exp \frac{\lambda}{t} \int_0^t \int_0^t Y_\varepsilon(W_s - W_\sigma) \, ds \, d\sigma \right] = 0.$$

PROOF. One can bound $\phi(x)$ which behaves like $\frac{1}{|x|}$ near 0 and like $\frac{1}{|x|^3}$ near ∞ by $\frac{C}{|x|^{3/2}}$. In particular,

$$Y_\varepsilon(x) \leq \frac{C\sqrt{\varepsilon}}{|x|^{3/2}}.$$

Then by time ordering and Jensen’s inequality,

$$\begin{aligned} & \exp \left\{ \frac{\lambda}{t} \int_0^t \int_0^t Y_\varepsilon(W_s - W_\sigma) \, ds \, d\sigma \right\} \\ &= \exp \left\{ \frac{2\lambda}{t} \int_0^t \left\{ \int_s^t Y_\varepsilon(W_s - W_\sigma) \, d\sigma \right\} ds \right\} \\ &\leq \frac{1}{t} \int_0^t \exp \left\{ 2\lambda \int_s^t Y_\varepsilon(W_s - W_\sigma) \, d\sigma \right\} ds \\ &\leq \frac{1}{t} \int_0^t \exp \left\{ 2C\lambda\sqrt{\varepsilon} \int_s^t \frac{1}{|W_s - W_\sigma|^{3/2}} \, d\sigma \right\} ds \\ &\stackrel{(D)}{=} \frac{1}{t} \int_0^t ds \exp \left\{ 2C\lambda\sqrt{\varepsilon} \int_0^{t-s} \frac{1}{|W_\sigma|^{3/2}} \, d\sigma \right\}. \end{aligned}$$

If we can show that, for $\varepsilon > 0$ small enough,

$$(51) \quad \sup_{x \in \mathbb{R}^3} \mathbb{E}^{(x)} \left\{ \exp \left\{ 2C\lambda\sqrt{\varepsilon} \int_0^1 \frac{1}{|W_\sigma|^{3/2}} \, d\sigma \right\} \right\} \leq \alpha < \infty,$$

then it follows, by successive conditioning and the Markov property,

$$(52) \quad \mathbb{E} \left\{ \exp \left\{ 2C\lambda\sqrt{\varepsilon} \int_0^{t-s} \frac{1}{|W_\sigma|^{3/2}} \, d\sigma \right\} \right\} \leq \alpha^{t-s}.$$

This will prove (50).

It remains to check (51). For this, we appeal to Portenko’s lemma (see [6]), which states that, if for a Markov process $\{\mathbb{P}^{(x)}\}$ and for a function $\tilde{V} \geq 0$,

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}^{(x)} \left\{ \int_0^1 \tilde{V}(W_s) \, ds \right\} \leq \eta < 1$$

then

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}^{(x)} \left\{ \exp \left\{ \int_0^1 \tilde{V}(W_s) \, ds \right\} \right\} \leq \frac{\eta}{1 - \eta} = \alpha < \infty.$$

Hence, to prove (51), we need to verify that

$$(53) \quad \sup_{x \in \mathbb{R}^3} \mathbb{E}^{(x)} \left\{ \int_0^1 \frac{d\sigma}{|W_\sigma|^{3/2}} \right\} \\ = \sup_{x \in \mathbb{R}^3} \int_{\mathbb{R}^3} dy \int_0^1 d\sigma \int_{\mathbb{R}^3} \frac{1}{|y|^{3/2}} \frac{1}{(2\pi\sigma)^{3/2}} \exp\left\{-\frac{(y-x)^2}{2\sigma}\right\} < \infty.$$

One can see that

$$\sup_{x \in \mathbb{R}^3} \int_{\mathbb{R}^3} dy \frac{1}{|y|^{3/2}} \frac{1}{(2\pi\sigma)^{3/2}} \exp\left\{-\frac{(y-x)^2}{2\sigma}\right\}$$

is attained at $x = 0$ because we can rewrite the integral by Parseval’s identity as

$$c \int_{\mathbb{R}^3} \exp\left\{-\frac{\sigma|\xi|^2}{2} + i\langle x, \xi \rangle\right\} \frac{1}{|\xi|^{3/2}} d\xi,$$

where $c > 0$ is a constant. When $x = 0$ the integral reduces to $\int_0^1 \sigma^{-3/4} d\sigma$ which is finite. \square

We continue with the proof of the Theorem 5.1. First, we prove a large deviation estimate for \widehat{Q}_t .

THEOREM 5.3. *For any closed set $F \subset \widetilde{\mathcal{X}}$*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \widehat{Q}_t[F] \leq - \inf_{\xi \in F} \widetilde{J}(\xi),$$

and for any open set $G \subset \widetilde{\mathcal{X}}$

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log \widehat{Q}_t[G] \geq - \inf_{\xi \in G} \widetilde{J}(\xi),$$

where, for $\xi = \{\tilde{\alpha}_j\} \in \widetilde{\mathcal{X}}$,

$$\widetilde{J}(\xi) = \tilde{\rho} - \sum_j \left\{ \int \frac{1}{|x-y|} \alpha_j(dx) \alpha_j(dy) - \widetilde{I}(\tilde{\alpha}_j) \right\}$$

and $\tilde{\rho}$ is given by

$$\tilde{\rho} = \sup_{\xi \in \widetilde{\mathcal{X}}} \sum_j \left\{ \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\psi_j^2(x) \psi_j^2(y)}{|x-y|} dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|_2^2 \right\}$$

and $\alpha_j(dx) = \psi_j^2(x) dx$ with $\sum_j \int_{\mathbb{R}^3} \psi_j^2(x) dx \leq 1$.

PROOF. We fix a closed set $F \subset \tilde{\mathcal{X}}$. Then, by definition,

$$(54) \quad \begin{aligned} \widehat{\mathbb{Q}}_t(F) &= \widehat{\mathbb{P}}_t(\tilde{L}_t \in F) \\ &= \frac{\mathbb{E}^{\mathbb{Q}_t}\{\exp\{(1/t) \int_0^t \int_0^t (1/|W_\sigma - W_s|) d\sigma ds\} \mathbb{1}_F\}}{\mathbb{E}^{\mathbb{Q}_t}\{\exp\{(1/t) \int_0^t \int_0^t (1/|W_\sigma - W_s|) d\sigma ds\}\}}, \end{aligned}$$

where \mathbb{Q}_t is the distribution of \tilde{L}_t in $\tilde{\mathcal{M}}_1(\mathbb{R}^d)$. We first handle the numerator. The denominator will be taken care of similarly. Recall the decomposition (49). Then with $\frac{1}{p} + \frac{1}{q} = 1$ and Hölder’s inequality, the numerator becomes

$$\begin{aligned} &\int_F \exp\left[\frac{1}{t} \int_0^t \int_0^t \{V_\varepsilon(|W_s - W_\sigma|) + Y_\varepsilon(|W_s - W_\sigma|)\} d\sigma ds\right] d\mathbb{Q}_t \\ &\leq \left[\int_F \exp\left\{\frac{1}{t} \int_0^t \int_0^t p V_\varepsilon(|W_s - W_\sigma|) d\sigma ds\right\} d\mathbb{Q}_t\right]^{1/p} \\ &\quad \times \left[\int_F \exp\left\{\frac{1}{t} \int_0^t \int_0^t q Y_\varepsilon(|W_s - W_\sigma|) d\sigma ds\right\} d\mathbb{Q}_t\right]^{1/q}. \end{aligned}$$

Taking logarithm, dividing by t , passing to $\limsup_{t \rightarrow \infty}$ and followed by $\varepsilon \rightarrow 0$,

$$\begin{aligned} &\limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_F \exp\left[\frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} d\sigma ds\right] d\mathbb{Q}_t \\ &\leq \limsup_{\varepsilon \rightarrow 0} \frac{1}{p} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_F \exp\left[\frac{1}{t} \int_0^t \int_0^t p V_\varepsilon(|W_s - W_\sigma|) d\sigma ds\right] d\mathbb{Q}_t \\ &\quad + \limsup_{\varepsilon \rightarrow 0} \frac{1}{q} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_F \exp\left[\frac{1}{t} \int_0^t \int_0^t q Y_\varepsilon(|W_s - W_\sigma|) d\sigma ds\right] d\mathbb{Q}_t. \end{aligned}$$

By Lemma 5.2, the second term is 0. For the first term, since for every $\varepsilon > 0$, $V_\varepsilon \in \mathcal{F}_2$, by Corollary 3.3, Proposition 4.4 and Varadhan’s lemma,

$$\begin{aligned} &\limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_F \exp\left[\frac{1}{t} \int_0^t \int_0^t p V_\varepsilon(|W_s - W_\sigma|) d\sigma ds\right] d\mathbb{Q}_t \\ &\leq \sup_{\xi \in F} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} p V_\varepsilon(x - y) \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right], \end{aligned}$$

where $\xi = \{\tilde{\alpha}_j\}$ and $\alpha_j(dx) = \psi_j^2(x) dx$ with $\sum_j \int_{\mathbb{R}^3} \psi_j^2(x) dx \leq 1$.

Since $V_\varepsilon(x) \leq \frac{1}{|x|}$ and $V_\varepsilon(x) \rightarrow \frac{1}{|x|}$ as $\varepsilon \rightarrow 0$, for any $p > 1$,

$$\begin{aligned} &\limsup_{\varepsilon \rightarrow 0} \sup_{\xi \in F} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} p V_\varepsilon(x - y) \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right] \\ &= \sup_{\xi \in F} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} p \frac{1}{|x - y|} \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right]. \end{aligned}$$

We can now let $p \rightarrow 1$ and obtain

$$\begin{aligned}
 (55) \quad & \limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_F \exp \left[\frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} d\sigma ds \right] d\mathbb{Q}_t \\
 & \leq \sup_{\xi \in F} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right].
 \end{aligned}$$

The lower bound

$$\begin{aligned}
 (56) \quad & \limsup_{t \rightarrow \infty} \frac{1}{t} \log \int_G \exp \left[\frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} d\sigma ds \right] d\mathbb{Q}_t \\
 & \geq \sup_{\xi \in G} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right],
 \end{aligned}$$

for open sets $G \subset \tilde{\mathcal{X}}$ follows immediately from Lemma 4.3. This derives the asymptotic behavior of the numerator in (54). For the denominator, we invoke (55) for $F = \tilde{\mathcal{X}}$ and (56) for $G = \tilde{\mathcal{X}}$ to deduce

$$\begin{aligned}
 (57) \quad & \lim_{t \rightarrow \infty} \frac{1}{t} \log \int_{\tilde{\mathcal{X}}} \exp \left[\frac{1}{t} \int_0^t \int_0^t \frac{1}{|W_s - W_\sigma|} d\sigma ds \right] d\mathbb{Q}_t \\
 & = \sup_{\xi \in \tilde{\mathcal{X}}} \left[\sum_j \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi_j^2(x) \psi_j^2(y) dx dy - \frac{1}{2} \sum_j \|\nabla \psi_j\|^2 \right] \\
 & = \tilde{\rho}.
 \end{aligned}$$

We apply (55), (56) and (57) to (54). The theorem is proved. \square

We need a lemma here to complete the proof of Theorem 5.1.

LEMMA 5.4. *The supremum in (57) is attained only when ξ consists of a single orbit $\tilde{\mu}$ with $\mu(dx) = \psi^2(x) dx$ for a unique radially symmetric ψ and $\int_{\mathbb{R}^3} \psi(x)^2 dx = 1$.*

PROOF. If we rescale with $\psi(x)$ being replaced by $\sigma^2 \psi(\sigma x)$, the expression

$$\sigma^8 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi^2(\sigma x) \psi^2(\sigma y) dx dy - \sigma^6 \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \psi(\sigma x)|^2 dx$$

becomes

$$\sigma^3 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi^2(x) \psi^2(y) dx dy - \frac{1}{2} \sigma^3 \int |\nabla \psi(x)|^2 dx$$

while the mass $\sigma^4 \int_{\mathbb{R}^3} \psi^2(\sigma x) dx$ becomes $\sigma \int_{\mathbb{R}^3} \psi^2(x) dx$. Therefore if we define

$$\rho(m) = \sup_{\int_{\mathbb{R}^3} h^2(x) dx = m} \left[\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x - y|} \psi^2(x) \psi^2(y) dx dy - \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \psi(x)|^2 dx \right]$$

then $\rho(m) = Cm^3$. In particular, $\rho(m_1 + m_2) > \rho(m_1) + \rho(m_2)$ proving that supremum in (57) is attained at a single orbit $\xi = \{\tilde{\mu}\}$ of total mass $\mu(\mathbb{R}^3) = 1$. According to Lieb's theorem (see [5]), the function ψ that maximizes

$$\left[\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{|x-y|} \psi^2(x) \psi^2(y) dx dy - \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \psi(x)|^2 dx \right]$$

subject to $\int_{\mathbb{R}^3} \psi^2(x) dx = 1$ is unique up to translation. \square

Acknowledgements. The first author would like to thank Erwin Bolthausen (Zurich) and Wolfgang Koenig (Berlin) for many helpful discussions. Both authors would like to thank three anonymous referees whose input led to a more elaborate version of the present manuscript.

REFERENCES

- [1] BOLTHAUSEN, E. and SCHMOCK, U. (1997). On self-attracting d -dimensional random walks. *Ann. Probab.* **25** 531–572. [MR1434118](#)
- [2] DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time. I–IV. *Comm. Pure Appl. Math.* **28** 1–47; **28** 279–301 (1975); **29** 389–461 (1979); **36** 183–212 (1983).
- [3] DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotics for the Wiener sausage. *Comm. Pure Appl. Math.* **28** 525–565. [MR0397901](#)
- [4] DONSKER, M. D. and VARADHAN, S. R. S. (1983). Asymptotics for the polaron. *Comm. Pure Appl. Math.* **36** 505–528. [MR0709647](#)
- [5] LIEB, E. H. (1976/77). Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation. *Stud. Appl. Math.* **57** 93–105. [MR0471785](#)
- [6] PORTENKO, N. I. (1975). Diffusion processes with an unbounded drift coefficient. *Teor. Veroyatnost. i Primenen.* **20** 29–39. [MR0375483](#)

COURANT INSTITUTE
OF MATHEMATICAL SCIENCES
AND WIAS BERLIN
251 MERCER STREET
NEW YORK, NEW YORK 10012
USA
E-MAIL: mukherjee@cims.nyu.edu

COURANT INSTITUTE
OF MATHEMATICAL SCIENCES
251 MERCER STREET
NEW YORK, NEW YORK 10012
USA
E-MAIL: varadhan@cims.nyu.edu