The Annals of Probability
2020, Vol. 48, No. 5, 2119–2144
https://doi.org/10.1214/19-AOP1392
© Institute of Mathematical Statistics, 2020

IDENTIFICATION OF THE POLARON MEASURE IN STRONG COUPLING AND THE PEKAR VARIATIONAL FORMULA

BY CHIRANJIB MUKHERJEE[1] AND S. R. S. VARADHAN[2]

[1] *Institut für Mathematische Stochastik, Fachbereich Mathematik und Informatik, Universität Münster,* *chiranjib.mukherjee@uni-muenster.de*

[2] *Courant Institute of Mathematical Sciences, New York University,* *varadhan@cims.nyu.edu*

The path measure corresponding to the *Fröhlich polaron* appearing in quantum statistical mechanics is defined as the tilted measure

$$\mathrm{d}\widehat{\mathbb{P}}_{\varepsilon,T} = \frac{1}{Z(\varepsilon,T)} \exp\left\{\frac{1}{2}\int_{-T}^{T}\int_{-T}^{T} \frac{\varepsilon \mathrm{e}^{-\varepsilon|t-s|}}{|\omega(t)-\omega(s)|}\,\mathrm{d}s\,\mathrm{d}t\right\}\mathrm{d}\mathbb{P}.$$

Here, $\varepsilon > 0$ is a constant known as the *Kac parameter* or the *inverse-coupling parameter*, and $\mathbb{P}$ is the distribution of the increments of the three-dimensional Brownian motion. In (*Comm. Pure Appl. Math.* **73** (2020) 350–383) it was shown that, when $\varepsilon > 0$ is sufficiently small or sufficiently large, the (thermodynamic) limit $\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T} = \widehat{\mathbb{P}}_{\varepsilon}$ exists as a process with stationary increments, and this limit was identified explicitly as a mixture of Gaussian processes. In the present article the *strong coupling limit* or the *vanishing Kac parameter* limit $\lim_{\varepsilon\to 0}\widehat{\mathbb{P}}_{\varepsilon}$ is investigated. It is shown that this limit exists and coincides with the *increments* of the so-called *Pekar process*, a stationary diffusion with generator $\frac{1}{2}\Delta + (\nabla\psi/\psi)\cdot\nabla$, where ψ is the unique (up to spatial translations) maximizer of the *Pekar variational problem*

$$g_0 = \sup_{\|\psi\|_2=1}\left\{\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \psi^2(x)\psi^2(y)|x-y|^{-1}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\|\nabla\psi\|_2^2\right\}.$$

As the Pekar process was also earlier shown (*Ann. Probab.* **44** (2016) 3934–3964; *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 2214–2228; *Comm. Pure Appl. Math.* **70** (2017) 1598–1629) to be the limiting object of the *mean-field polaron measures*, the present identification of the strong coupling limit is a rigorous justification of the mean-field approximation of the polaron problem (on the level of path measures) conjectured by Spohn in (*Ann. Physics* **175** (1987) 278–318). Replacing the Coulomb potential by continuous function vanishing at infinity and assuming uniqueness (modulo translations) of the relevant variational problem, our proof also shows that path measures coming from a Kac interaction of the above form with translation invariance in space converge to the increments of the corresponding mean-field model.

1. Introduction and summary.

Consider any function $V:\mathbb{R}^d\to\mathbb{R}$ which is rotationally symmetric, continuous and vanishes at infinity. Alternatively, choose the Coulomb potential $V(x) = \frac{1}{|x|}$ in $\mathbb{R}^3$ which explodes at the origin. For any $T > 0$, let $\mathbb{P} = \mathbb{P}_T$ be the law of Brownian increments in the time interval $[-T, T]$ (i.e., $\mathbb{P}$ is defined only on the σ-algebra generated by $\{\omega(t)-\omega(s)\colon -T\le s<t\le T\}$), and consider a transformed path measure which descends from a *Kac interaction* of the form

$$\mathrm{d}\widehat{\mathbb{P}}_{\varepsilon,T} = \frac{1}{Z(\varepsilon,T)} \exp\left\{\frac{1}{2}\int_{-T}^{T}\int_{-T}^{T} \varepsilon\mathrm{e}^{-\varepsilon|t-s|}V\big(\omega(t)-\omega(s)\big)\,\mathrm{d}s\,\mathrm{d}t\right\}\mathrm{d}\mathbb{P}, \tag{1}$$

Received December 2018.
MSC2020 subject classifications. Primary 60J65, 60J35; secondary 60F10.
Key words and phrases. Fröhlich polaron, strong coupling limit, Kac interaction, Pekar variational formula, Pekar process, Landau–Pekar theory.

where $\varepsilon > 0$ is a constant, or the *Kac-parameter*. Here, $Z_{\varepsilon,T}$ is the total mass of the exponential weight above, or the *partition function*, whose asymptotic behavior in the thermodynamic limit $T \to \infty$, followed by the vanishing Kac-parameter limit $\varepsilon \to 0$, has been investigated in [3], where it is shown that

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\frac{1}{T}\log Z_{\varepsilon,T} = \sup_{\substack{\psi\in H^1(\mathbb{R}^d)\\ \|\psi\|_2=1}}\left\{\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\psi^2(x)\psi^2(y)V(x-y)\,\mathrm{d}x\,\mathrm{d}y-\frac{1}{2}\|\nabla\psi\|_2^2\right\}. \tag{2}$$

Assuming that the above variational problem admits a maximizer ψ, which is unique modulo spatial translations, the main result of the article shows that the actual path measures $\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}$ converge in the vanishing Kac interaction to the increments of a stationary diffusion process with generator $\frac{1}{2}\Delta + (\nabla\psi/\psi)\cdot\nabla$. The incentive of our present work came from proving the above convergence of the path measures for the particular choice $V(x) = \frac{1}{|x|}$ in $d = 3$, for which the above variational problem is known to admit a smooth maximizer, which is unique, up to translations [8]. This choice corresponds to the *Fröhlich Polaron*—a model which enjoys quite some prominence in quantum statistical mechanics.

1.1. *Fröhlich polaron and its path measures.* Physical motivation of the Fröhlich polaron originates from studying the effective behavior of a slow moving electron coupled to a crystal. For the physical relevance of this model, we refer to [5, 7, 14, 15]. The probabilistic layout of this problem was also founded by Feynman, via the *path integral* formulation, which is captured by studying the behavior of a Gibbs measure supported on a three-dimensional Brownian motion acting under a selfattractive Coulomb interaction; see also [13], Section 1, for a discussion which relates the quantum mechanical background of the polaron to the path integral approach, which we pursued in [13] and continue to do so in the present context. This selfattractive interaction then defines the tilted measure of the form

$$\widehat{\mathbb{P}}_{\varepsilon,T}(\mathrm{d}\omega) = \frac{1}{Z(\varepsilon,T)}\exp\left\{\frac{\varepsilon}{2}\int_{-T}^{T}\int_{-T}^{T}\frac{\mathrm{e}^{-\varepsilon|t-s|}}{|\omega(t)-\omega(s)|}\,\mathrm{d}\sigma\,\mathrm{d}s\right\}\mathbb{P}(\mathrm{d}\omega). \tag{3}$$

In the above display $\varepsilon > 0$ is a constant, and, if we set

$$\varepsilon = \alpha^{-2}, \tag{4}$$

α is called the *coupling parameter*. As before and also in (3), $\mathbb{P}$ refers to the law of three-dimensional white noise, which is defined only on the σ-field generated by three-dimensional Brownian increments $\{\omega(t)-\omega(s)\colon -T \le s < t \le T\}$, while $Z(\varepsilon,T)$ is the normalization constant or the *partition function*.

Here, we are concerned with the physically relevant regime of the *strong coupling limit* initiated already by Pekar [14]. This regime corresponds to studying the asymptotic behavior of the interaction (3) as $T \to \infty$, followed by $\varepsilon \to 0$ (or $\alpha \to \infty$). Note that, for any $\varepsilon > 0$, replacing $\omega(s)$ by $\sqrt{\varepsilon}\omega(\frac{s}{\varepsilon})$ and invoking the scaling property of Brownian motion, we get

$$\begin{aligned}
Z(\varepsilon,T) &= \mathbb{E}^{\mathbb{P}}\left[\exp\left\{\frac{\varepsilon}{2}\int_{-T}^{T}\int_{-T}^{T}\frac{\mathrm{e}^{-\varepsilon|t-s|}\,\mathrm{d}s\,\mathrm{d}t}{|\omega(t)-\omega(s)|}\right\}\right]\\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left\{\frac{1}{2\sqrt{\varepsilon}}\int_{-\varepsilon T}^{\varepsilon T}\int_{-\varepsilon T}^{\varepsilon T}\frac{\mathrm{e}^{-|t-s|}\,\mathrm{d}s\,\mathrm{d}t}{|\omega(t)-\omega(s)|}\right\}\right]\\
&= Z\left(\frac{1}{\sqrt{\varepsilon}},\varepsilon T\right)\\
&= Z\left(\alpha,\frac{T}{\alpha^2}\right),
\end{aligned}$$

where the last identity follows from (4). It was conjectured in [14] that the *ground-state energy* of the strong coupling polaron

$$
(5)\qquad \begin{aligned} g_0 &\overset{\text{(def)}}{=} \lim_{\varepsilon\to 0}\lim_{T\to\infty}\frac{1}{2T}\log Z(\varepsilon,T)=\lim_{\alpha\to\infty}\lim_{T\to\infty}\frac{1}{2T}\log Z\big(\alpha,\alpha^{-2}T\big)\\ &=\lim_{\alpha\to\infty}\frac{1}{\alpha^2}\lim_{T\to\infty}\frac{1}{2T}\log Z(\alpha,T)\end{aligned}
$$

exists and is given by the *Pekar variational formula*

$$
(6)\qquad g_0=\sup_{\substack{\psi\in H^1(\mathbb{R}^3)\\ \|\psi\|_2=1}}\bigg\{\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\mathrm{d}x\,\mathrm{d}y\,\frac{\psi^2(x)\psi^2(y)}{|x-y|}-\frac{1}{2}\|\nabla\psi\|_2^2\bigg\}.
$$

Here, $H^1(\mathbb{R}^3)$ denotes the usual Sobolev space of square integrable functions with square integrable gradient. Pekar's conjecture was proved in [3] using large deviation theory (see also [10] for a different approach), and the Pekar variational formula was analyzed by Lieb [8] who showed that the supremum in (6) is attained and the maximizing set consists of only translates

$$
\mathfrak{m}=\big\{\psi_y(\cdot)=\psi_0(\cdot-y):y\in R^3\big\}
$$

of a single maximizer ψ_0 which is rotationally symmetric around 0.

Apart from the ground state energy, another relevant physical quantity for the polaron is its so-called *effective mass*, whose rigorous definition requires investigating the (asymptotic) behavior of the polaron path measures $\widehat{\mathbb{P}}_{\varepsilon,T}$. Unlike the partition function, a rigorous analysis of the actual path measures $\widehat{\mathbb{P}}_{\varepsilon,T}$ turned out to be much more subtle and had remained unanswered on a rigorous level. In a recent article [13] we have shown that there exists $\varepsilon_0,\varepsilon_1\in(0,\infty)$ such that for any fixed $\varepsilon\in(0,\varepsilon_0)\cup(\varepsilon_1,\infty)$, the limit

$$
\widehat{\mathbb{P}}_\varepsilon=\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}
$$

exists and identified the limit $\widehat{\mathbb{P}}_\varepsilon$, *explicitly*. As a corollary we have also deduced the central limit theorem for the distributions

$$
(7)\qquad \widehat{\mathbb{P}}_{\varepsilon,T}\bigg(\frac{1}{\sqrt{2T}}\big[\omega(T)-\omega(-T)\big]\bigg)^{-1}\Rightarrow\mathbf{N}\big(0,\sigma^2(\varepsilon)\mathbf{Id}\big)
$$

of the increment of the process under $\widehat{\mathbb{P}}_{\varepsilon,T}$ and obtained an explicit formula for the limiting variance $\sigma^2(\varepsilon)\in(0,1)$ which is directly related to the aforementioned effective mass $m_{\mathrm{eff}}(\varepsilon)$ of the polaron.[1] It is the goal of the present article to investigate and characterize the strong coupling limit

$$
\widehat{\mathbb{Q}}=\lim_{\varepsilon\to 0}\widehat{\mathbb{P}}_\varepsilon=\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}
$$

of the polaron measures. As we will see, this limit $\widehat{\mathbb{Q}}$ will be determined *uniquely* by *any* maximizer ψ of the Pekar variational problem (6).

[1] The relation $m_{\mathrm{eff}}(\varepsilon)^{-1}=\sigma^2(\varepsilon)$ follows as a direct consequence of our CLT result (7); see [4]. It has also been shown recently in [9] that $\lim_{\varepsilon\to 0}m_{\mathrm{eff}}(\varepsilon)=\infty$. However, the rate of divergence of the latter quantity is not known.

1.2. *The mean-field Fröhlich polaron and the Pekar process.* Before turning to a more formal description of our main results, it is useful to provide an intuitive interpretation of (5) and (6). We remark that the interaction appearing in the polaron problem (3) is *selfattractive*. For fixed $\varepsilon > 0$, the measure $\widehat{\mathbb{P}}_{\varepsilon,T}$ favors paths which make $|\omega(t)-\omega(s)|$ small, when $|t-s|$ is not large. In other words, these paths tend to clump together on short time scales. However, for strong coupling this interaction becomes more and more smeared out, and on an intuitive level in this regime (i.e., $\varepsilon \downarrow 0$) one expects the polaron interaction to resemble the *mean-field interaction* given by

$$\widehat{\mathbb{P}}_T^{(\mathrm{mf})}(\mathrm{d}\omega) = \frac{1}{Z_T^{(\mathrm{mf})}} \exp\biggl\{\frac{1}{T}\int_0^T\int_0^T \mathrm{d}t\,\mathrm{d}s \frac{1}{|\omega(t)-\omega(s)|}\biggr\}\mathbb{P}(\omega). \tag{8}$$

The earlier result (5) indeed justified this intuition and underlined the parallel behavior for the partition functions (on a logarithmic scale) of these two models,

$$g_0 = \lim_{\varepsilon\to 0}\lim_{T\to\infty}\frac{1}{T}\log Z(\varepsilon,T) = \lim_{T\to\infty}\frac{1}{T}\log Z_T^{(\mathrm{mf})}.$$

Based on the above intuition, Spohn [15] conjectured that the the strong coupling behavior of the actual polaron measures $\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{Q}}_{\varepsilon,T}$ should be closely related to the behavior of its mean-field counterpart $\lim_{T\to\infty}\widehat{\mathbb{P}}_T^{(\mathrm{mf})}$. Assuming this conjecture to be true, he also heuristically derived ([15]) the actual decay rate (in leading order) of the diffusion constant $\sigma^2(\varepsilon)$ of the central limit theorem appearing in (7) as $\varepsilon \to 0$.

A rigorous analysis of the mean-field model (8) was determined in [1], based on the theory developed in [12], and its extension [6]. It was shown in [1] that the distribution $\widehat{\mathbb{P}}_T^{(\mathrm{mf})}L_T^{-1}$ of the Brownian occupation measures $L_T = \frac{1}{T}\int_0^T \delta_{\omega_s}\,\mathrm{d}s$ under the mean-field model converges to the distribution of a random translation $[\psi_0^2 \star \delta_X]\,\mathrm{d}z$ of $\psi_0^2\,\mathrm{d}z$, with the random shift X having a density $c_0\psi_0$, where ψ_0 is the maximizer in (6) centered at 0 and c_0 is the normalizing constant. Furthermore, it was also shown ([1]) that the mean-field measures themselves converge

$$\widehat{\mathbb{P}}_T^{(\mathrm{mf})} \Rightarrow c_0\int_{\mathbb{R}^3}\mathbb{Q}_{\psi_y}\psi_0(y)\,\mathrm{d}y \tag{9}$$

to a mixture with weight $c_0\psi_0(y)\,\mathrm{d}y$ of the diffusion processes $\mathbb{Q}_{\psi_y}$ with generator

$$\frac{1}{2}\Delta + \frac{\nabla\psi_y}{\psi_y}\cdot\nabla \tag{10}$$

initialized to start from 0. The heuristic definition of this diffusion process was set forth in [15] and was called the *Pekar process*.

1.3. *Main results*: *Strong coupling/vanishing Kac limit toward increments of Pekar process.* Note that the distribution of the *increments* of the stationary versions of the Pekar process $\mathbb{Q}_{\psi_y}$ with generator (10) does not depend on y and defines a unique process $\widehat{\mathbb{Q}} = \widehat{\mathbb{Q}}_\psi$ on the space of increments. In the present context our main result is stated as follows:

THEOREM 1.1 (Convergence of the Fröhlich polaron measure in strong coupling). *Let $\widehat{\mathbb{P}}_{\varepsilon,T}$ be the path measures for the Fröhlich polaron defined in* (3), *and let $\widehat{\mathbb{Q}}_\psi$ be the common distribution of the increments of the stationary Pekar process $\mathbb{Q}_{\psi_y}$ with generator* (10). *Then,*

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}(\cdot) = \lim_{\varepsilon\to 0}\widehat{\mathbb{P}}_\varepsilon(\cdot) = \widehat{\mathbb{Q}}_\psi(\cdot).$$

Thus, the above result justifies the conjecture posed in [15] regarding the parallel behavior of the measures $\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}$ and $\lim_{T\to\infty}\widehat{\mathbb{P}}^{(\mathrm{mf})}_T$.

While we are intrinsically interested in analyzing the case of the Fröhlich polaron, the method of our proof is robust enough to show quite generally that path measures coming from any Kac interaction of the form (1) with translation invariance in space converge to the increments of the corresponding mean-field model. More precisely, we have, as our second main result,

THEOREM 1.2 (Convergence of general Kac interactions toward increments of Pekar-type process). *Let $V:\mathbb{R}^d\to\mathbb{R}$ be any rotationally symmetric and continuous function vanisihing at infinity, and assume that the corresponding variational problem* (2) *admits a smooth, strictly positive maximizer $\psi^{(V)}\in H^1(\mathbb{R}^d)$ which is unique modulo spatial translations. Then, we have the following identification of the path measures defined in* (1) *in the vanishing Kac limit*:

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{\mathbb{P}}^{(V)}_{\varepsilon,T}(\cdot)=\widehat{\mathbb{Q}}_{\psi^{(V)}}(\cdot),$$

where $\widehat{\mathbb{Q}}_{\psi^{(V)}}$ is the common distribution of the increments of the stationary process with generator $\frac{1}{2}\Delta+\frac{\nabla\psi^{(V)}}{\psi^{(V)}}\cdot\nabla$.

Let us first briefly outline the method developed here for proving the above results.

1.4. *Outline of the proof.* In order to provide some guidelines for the reader, we will conclude the Introduction with an outline of the proof of our main result.

Let Ω_0 denote the space of continuous functions on $\omega:\mathbb{R}\to\mathbb{R}^d$ vanishing at the origin. Then for any $t\in\mathbb{R}$, we have a shift $\theta_t:\Omega_0\to\Omega_0$ defined via $(\theta_t\omega)(\cdot)=\omega(t+\cdot)-\omega(\cdot)$, and we can denote by $\mathcal{M}_{\mathrm{si}}(\Omega_0)$ the space of θ_t-invariant probability measures on Ω_0, or the space of *processes with stationary increments*. Note that we also have an action $\theta_t:\Omega_0\otimes\mathbb{R}^d\to\Omega_0\otimes\mathbb{R}^d$ by $\theta_t(\omega,x)=(\omega(t+\cdot)-\omega(t),x+\omega(t))$. Then, we can denote by $\mathcal{M}_{\mathrm{s}}(\Omega_0\otimes\mathbb{R}^d)$ to be the space of θ_t-invariant probability measures on $\Omega_0\otimes\mathbb{R}^d$, or the space of *stationary processes*.

The first main step for our proof is to show that $\widehat{\mathbb{P}}_\varepsilon\in\mathfrak{m}_\varepsilon$, where $\mathfrak{m}_\varepsilon$ is the set of maximizers of the variational problem (for the particular case $d=3$)

$$g(\varepsilon)=\sup_{\mathbb{Q}\in\mathcal{M}_{\mathrm{si}}(\Omega_0)}\biggl[\mathbb{E}^{\mathbb{Q}}\biggl(\int_0^\infty\frac{\varepsilon e^{-\varepsilon t}\,dt}{|\omega(t)-\omega(0)|}\biggr)-H(\mathbb{Q}|\mathbb{P})\biggr]. \tag{11}$$

As mentioned earlier, a variational formula of the above form for $\lim_{T\to\infty}\frac{1}{T}\log Z_{\varepsilon,T}$ was first obtained in [3] where the supremum above was taken over all stationary processes in $\mathcal{M}_{\mathrm{s}}(\Omega_0\otimes\mathbb{R}^3)$. This result was a consequence of a weak large deviation principle (LDP) for the empirical process of Brownian motion. However, in this case, the supremum may not be attained. This issue is resolved if we exploit the underlying i.i.d. structure of the noise which provides exponential tightness and a full LDP for the emprical process of Brownian increments. In this set-up uniform relative entropy estimates then show that the variational formula (11) is coercive which guarantees the existence of at least one maximizer in $\mathfrak{m}_\varepsilon$, and, moreover, $\bigcup_{\varepsilon<\varepsilon_0}\mathfrak{m}_\varepsilon$ is also tight. The above strong LDP, combined with the existence of the actual limit $\lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}=\widehat{\mathbb{P}}_\varepsilon$, then also shows that $\widehat{\mathbb{P}}_\varepsilon\in\mathfrak{m}_\varepsilon$.

It remains to show that if $\varepsilon_n\to 0$ and $(\mathbb{Q}_n)\subset\mathfrak{m}_{\varepsilon_n}$ is any sequence of maximimizers such that $\mathbb{Q}_n\Rightarrow\mathbb{Q}$ weakly, then $\mathbb{Q}\in\mathcal{M}_{\mathrm{si}}(\Omega_0)$ must be the distribution of the increments of the stationary Pekar process. The task then splits into two further steps. First, note that not every process with stationary increments appear as the *increments* of another stationary process. For our purposes we first provide a general criterion that determines when any $\mathbb{Q} \in \mathcal{M}_{\mathrm{si}}(\Omega_0)$ admits this cocycle representation (i.e., any $\mathbb{Q} \in \mathcal{M}_{\mathrm{si}}(\Omega_0)$ appears as the increments of some $\mathbb{Q}' \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$); see Theorem 3.1. This criterion is formulated in terms of convergence of integrals of continuous functions vanishing at infinity w.r.t. measures on the function space Ω_0. However, since Ω_0 is not even locally compact, there is no notion of usual vague convergence of measures on this space (determined by convergence of integrals w.r.t. continuous functions vanishing at infinity). We therefore formulate a notion of *wea-gue* convergence on measures on $\Omega_0 \otimes \mathbb{R}^d$; see Section 2.1 which is conceptually important for the proof of Theorem 3.1.

To this end, however, we encounter another fundamental problem. The cocycle representation obtained from Theorem 3.1 is not unique; any $\mathbb{Q} \in \mathcal{M}_{\mathrm{si}}(\Omega_0)$ can be written as the increments of an *entire orbit* $\widetilde{\mathbb{Q}}' = \{\mathbb{Q} \star \delta_a \colon a \in \mathbb{R}^d\}$ for some $\mathbb{Q}' \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$, where for any $A \subset \Omega_0$ and $B \subset \mathbb{R}^d$, we define $(\mathbb{Q}' \star \delta_a)[A \otimes B] = \mathbb{Q}'[A \otimes (B - a)]$. While identifying any limiting maximizer as the increments of the Pekar process, this nonuniqueness leads to the following obstacle. For any sequence $(\mathbb{Q}_n)_n \subset \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$, even if we assume that its marginals $\mathbb{Q}_n^{(1)}$ on Ω_0 form a uniformly tight family, its marginals $\mathbb{Q}_n^{(2)}$ on $\mathbb{R}^d$ might still fail to have a convergent subsequence. Its mass may split and escape into two or more different directions (for instance, $\sum_i p_i \delta_{a_n^{(j)}}$ such that $|a_n^{(i)} - a_n^{(j)}| \to \infty$ for $i \neq j$); it could totally disintegrate into dust like a Gaussian with large variance, or it could form a mixture of all these *widely separated* components. By taking spatial shifts and recovering one such component at a time, in the limit we only imagine an empty, finite or countable collection of orbits of subprobability meaures, while possibly allowing some mass to totally disintegrate into dust. This intuition leads to a refinement of the method developed in [12]. We define the space $\mathscr{X} = \{\Theta = [\xi, \beta]\}$ of all collections $[\xi, \beta]$ where $\xi = \{\widetilde{\lambda}_j\}$ is a empty, finite or countable collection of orbits of subprobability measures on $\Omega_0 \otimes \mathbb{R}^d$ and β is a probability measure on $\mathbb{R}^d$ such that $\sum_j \widetilde{\lambda}^{(1)}(\cdot) \leq \beta(\cdot)$, where $\widetilde{\lambda}^{(1)}$ is the common marginal of λ on Ω_0; see Section 2.2. We metrize the space $\mathscr{X}$ which provides a topology that ensures the following crucial property: Let $K \subset \mathscr{X}$ be such that, as $\Theta = [\xi, \beta]$ varies over K, β varies over a uniformly tight family of probability measures on Ω_0. Then, every sequence in K finds a convergent subsequence in that metric in $\mathscr{X}$. On a technical level the above recipe then leads to a generalization of the theory developed in [12] where the compactification was earlier carried out for measures only in $\mathbb{R}^d$. In the present context this refined compactness result for the space $\mathscr{X}$ holds the key for the identification of any limiting maximizer $\lim_{\varepsilon \to 0} \widehat{\mathbb{P}}_\varepsilon$ as the distribution of increments of the stationary Pekar process.

Organization of the article: The rest of the article is organized as follows. Section 2 is entirely devoted to defining topologies on measures on suitable spaces, their quotient spaces under group actions and deriving properties of the aforementioned space $\mathscr{X}$. We will derive several useful properties regarding processes with stationary increments and their cocycle representations in Section 3. Section 4 is devoted to proving relative entropy estimates that provide coercivity properties of the variational formula $g(\varepsilon)$, while enabling us to circumvent singularity of the Coulomb potential. Combining all previous arguments, in Section 4.4 and Section 5 we prove our key results showing that $\widehat{\mathbb{P}}_\varepsilon \in \mathfrak{m}_\varepsilon$, while as $\varepsilon \to 0$, any limiting maximizer in $\mathfrak{m}_\varepsilon$ appears as the distribution of increments of the Pekar process.

2. Compactification of quotient spaces under group actions.

2.1. *The wea-gue topology.* In the sequel, for any topological space Y we will write for $\mathcal{M}_1(Y)$ and $\mathcal{M}_{\leq 1}(Y)$ to be the spaces of all probability and subprobability measures on Y, respectively. We now fix a complete separable metric space X. Then, both $\mathcal{M}_1(X \otimes$

$\mathbb{R}^d$) and $\mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$ are equipped with the weak topology for which a sequence λ_n of (sub)probability measures converge to λ, written $\lambda_n \Rightarrow \lambda$, if and only if

$$\int F(x,y)\lambda_n(\mathrm{d}x\,\mathrm{d}y) \to \int F(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y) \tag{12}$$

for any continuous and bounded function $F : X \otimes \mathbb{R}^d \to \mathbb{R}$. The same notion of weak convergence holds also for the space $\mathcal{M}_{\le 1}(\mathbb{R}^d)$. However, for the latter case we also have the notion of *vague convergence* (written $\lambda_n \hookrightarrow \lambda$) which demands (12) to hold for continuous functions $f : \mathbb{R}^d \to \mathbb{R}$ vanishing at infinity or for continuous functions with compact support. Equivalently, the vague convergence can be also obtained by considering the weak topology on (sub-)probability measures on the one-point compactification $\overline{\mathbb{R}}^d = \mathbb{R}^d \cup \{\infty\}$ and removing the mass at ∞. Note that any sequence of subprobability measures on $\mathbb{R}^d$ has a vaguely convergent subsequence, while the weak convergence fails to possess this property.

Since the space X need not be locally compact, there is no notion of vague convergence for measures on $X \otimes \mathbb{R}^d$. However, we can again consider the weak topology on $\mathcal{M}_{\le 1}(X \otimes \overline{\mathbb{R}}^d)$ and remove the mass at ∞, which leads us to the following notion of *wea-gue* convergence in this set-up: We say that a sequence λ_n converges to λ *weaguely* in the space $\mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$ if and only if $\int_{X\otimes\mathbb{R}^d} F(x,y)\lambda_n(\mathrm{d}x\,\mathrm{d}y) \to \int_{X\otimes\mathbb{R}^d} F(x,y)\lambda(\mathrm{d}x\,\mathrm{d}y)$ for all continuous functions $F : X \otimes \mathbb{R}^d \to \mathbb{R}$ such that

$$\lim_{|y|\to\infty} \sup_{x\in X} F(x,y) = 0.$$

We will have several occasions to use the following elementary result which is an immediate consequence of the aforementioned compactness of vague topology and Prohorov's theorem.

LEMMA 2.1. *Any sequence of probability measures on $X \otimes R^d$ with uniformly tight marginals on X will have a subsequence that converges weaguely.*

2.2. *Quotient space of $\mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$ and the metric space $(\mathscr{X}, \mathscr{D})$.* Note that the translation group $\{T_a : a \in \mathbb{R}^d\}$ acts on $X \otimes \mathbb{R}^d$ by mapping

$$(x,y) \mapsto (x, y+a). \tag{13}$$

This action then also induces a map on the space of measures $\mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$, which we denote by

$$\lambda(\mathrm{d}x\,\mathrm{d}y) \mapsto (\lambda \star \delta_a)(\mathrm{d}x\,\mathrm{d}y) \stackrel{\text{(def)}}{=} \lambda\big(\mathrm{d}x\mathrm{d}(y+a)\big), \quad x \in X, a, y \in \mathbb{R}^d.$$

We can then define an equivalence relation on $\mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$ by setting

$$\lambda \sim \lambda' \quad \text{if } \lambda' = \lambda \star \delta_a \text{ for some } a \in \mathbb{R}^d,$$

which leads to the notion of equivalence classes or *orbits* which we denote by $\widetilde{\lambda} = \{\lambda \star \delta_a : a \in \mathbb{R}^d\}$ and the corresponding quotient space by

$$\widetilde{\mathcal{M}}_{\le 1}(X \otimes \mathbb{R}^d) = \mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)/\sim.$$

For any $\lambda \in \mathcal{M}_{\le 1}(X \otimes \mathbb{R}^d)$, $\lambda^{(1)}$ and $\lambda^{(2)}$ are its marginals on X and $\mathbb{R}^d$, respectively. If $\widetilde{\lambda} \in \widetilde{\mathcal{M}}_{\le 1}(X \otimes \mathbb{R}^d)$ is an orbit, $\widetilde{\lambda}^{(1)}$ is well defined as the common marginal, whereas $\widetilde{\lambda}^{(2)}$ is an orbit of subprobability measures on $\mathbb{R}^d$.

We enlarge the space $\widetilde{\mathcal{M}}_{\le 1}(X \otimes \mathbb{R}^d)$ to

$$\mathscr{X} = \Big\{\Theta = [\xi, \beta] : \xi = \{\widetilde{\lambda}_j\}_j, \lambda_j \in \mathcal{M}_{\le 1}\big(X \otimes \mathbb{R}^d\big), \beta \in \mathcal{M}_1(X), \sum_{\widetilde{\lambda}\in\xi} \widetilde{\lambda}^{(1)}(\cdot) \le \beta(\cdot)\Big\}. \tag{14}$$

In other words, $\mathscr{X}$ consists of all collections of $\Theta = [\xi, \beta]$, where β is a probability measure on X and ξ is an empty, finite or countable collection $\{\widetilde{\lambda}_j\}$ of orbits $\widetilde{\lambda}_j \in \widetilde{\mathcal{M}}_{\le 1}(X \otimes \mathbb{R}^d)$ with the property that

$$\sum_{\widetilde{\lambda} \in \xi} \widetilde{\lambda}^{(1)}(\cdot) \le \beta(\cdot).$$

Clearly, we have an embedding

$$\widetilde{\mathcal{M}}_1(X \otimes \mathbb{R}^d) \hookrightarrow \mathscr{X},$$

since for any single orbit $\widetilde{\lambda}$ of a probability measure, we have

$$[\{\widetilde{\lambda}\}, \widetilde{\lambda}^{(1)}] \in \mathscr{X}.$$

We will now define a metric on $\mathscr{X}$. For any $\Theta_1 = [\xi_1, \beta_1]$, $\theta_2 = [\xi_2, \beta_2] \in \mathscr{X}$, we set

$$\mathscr{D}(\Theta_1, \Theta_2) = \mathbf{D}^\star(\xi_1, \xi_2) + \mathrm{d}(\beta_1, \beta_2), \tag{15}$$

where $\mathbf{D}^\star$ and d are defined as follows. The definition of d as a metric on $\mathcal{M}_1(X)$ is straightforward. Indeed, we choose a countable set of continuous functions $\mathfrak{F} = \{f_j(x)\}$ with $\sup_j \sup_x |f_j(x)| \le 1$ such that the existence of the limit $\lim_{n\to\infty} |\int f_j(x)\beta_n(\mathrm{d}x) - \int f_j(x)\beta(\mathrm{d}x)| = 0$ for every j is equivalent to the weak convergence of $\beta_n \Rightarrow \beta$. On $\mathcal{M}_1(X)$, we then have the metric

$$\mathrm{d}(\beta_1, \beta_2) = \sum_j \frac{1}{2^j} \left| \int_X f_j(x)\beta_1(\mathrm{d}x) - \int_X f_j(x)\beta_2(\mathrm{d}x) \right|. \tag{16}$$

We define $\mathbf{D}^\star$ as follows. For each $k \ge 2$, we denote by $\mathfrak{C}_k$ the space of functions $W(y_1, \dots, y_k)$ that satisfy $W(y_1 + a, \dots, y_k + a) = W(y_1, \dots, y_k)$ for all $a \in \mathbb{R}^d$ and $\lim_{\sup_{i,j} |y_i - y_j| \to \infty} |W(y_1, \dots, y_k)| = 0$. Since $\mathfrak{C}_k$ is separable in the uniform metric, we can choose a countable collection $\mathcal{W}_k \subset \mathfrak{C}_k$ of functions W such that they are uniformly bounded by 1, and their linear combinations are dense in $\mathfrak{C}_k$. We denote by $\mathcal{W}$ the countable set $\bigcup_{k\ge 2} \mathcal{W}_k$ and list it as $\{W_j\}$. Then, with $\mathfrak{F} = \{f_j\}_j$ being the basis for the metric d on $\mathcal{M}_1(X)$ above, we can enumerate all the combinations $\{f_j, W_{j'}\}$ as single sequence $\{f_r, W_r\}$ so that each W_r is a function $W(y_1, \dots, y_{k(r)})$ of $k(r)$ variables. Then, for any $\xi = \{\widetilde{\lambda}_j\}$, we set

$$\Lambda_{f,W}(\xi) = \sum_{\widetilde{\lambda} \in \xi_i} \int f(x_1) \cdots f(x_{k(r)}) W(y_1, \dots, y_{k(r)}) \prod_{i=1}^{k_r} \lambda(\mathrm{d}x_i, \mathrm{d}y_i) \tag{17}$$

and define

$$\mathbf{D}^\star(\xi_1, \xi_2) = \sum_{r=1}^{\infty} \frac{1}{2^r} \left| \Lambda_{f_r, W_r}(\xi_1) - \Lambda_{f_r, W_r}(\xi_2) \right|. \tag{18}$$

Note that the integrals in (17) depend only on the orbit $\widetilde{\lambda}$ and, therefore, $\mathbf{D}^\star$ is well defined. We now need to justify that $\mathscr{D}(\theta_1, \theta_2) = \mathbf{D}^\star(\xi_1, \xi_2) + \mathrm{d}(\beta_1, \beta_2)$, defined in (15), is a metric on $\mathscr{X}$. Since nonnegativity and triangle inequality is obvious, we only need to verify that $\mathscr{D}(\Theta_1, \Theta_2) = 0$ implies $\Theta_1 = \Theta_2$. However, since d is already a metric in $\mathcal{M}_1(X)$, we only need to verify that $\mathbf{D}^\star(\xi_1, \xi_2) = 0$ forces $\xi_1 = \xi_2$. The following lemma will guarantee the validity of the last statement:

LEMMA 2.2. *Let λ, γ be two probability measures on $X \otimes \mathbb{R}^d$ such that for any $k \ge 2$,*

$$\begin{aligned} &\int W(y_1, y_2, \dots, y_k) \prod_{i=1}^k f(x_i) \prod_{i=1}^k \lambda(\mathrm{d}x_i, \mathrm{d}y_i) \\ &\quad = \int W(y_1, y_2, \dots, y_k) \prod_{i=1}^k f(x_i) \prod_{i=1}^k \gamma(\mathrm{d}x_i, \mathrm{d}y_i) \end{aligned} \tag{19}$$

for all functions $f \in \mathfrak{F}$ and $W \in \mathcal{W}$. Then, there is $a \in \mathbb{R}^d$ such that $\gamma(A \times B) = \lambda(A \times (B + a))$ for all measurable $A \subset X$ and $B \subset \mathbb{R}^d$.

PROOF. Given (19), if we let $W \to 1$, the bounded convergence theorem implies that the marginals of λ and γ on X are the same. For any $f \geq 0$, let us define two measures λ_f and γ_f on $\mathbb{R}^d$ by

$$\lambda_f(B) = \frac{1}{c(f)} \int_{X \otimes B} f(x)\lambda(\mathrm{d}x, \mathrm{d}y), \qquad \gamma_f(B) = \frac{1}{c(f)} \int_{X \otimes B} f(x)\gamma(\mathrm{d}x, \mathrm{d}y),$$

where $c(f)$ is the normalizing constant

$$c(f) = \int_{X \otimes \mathbb{R}^d} f(x)\lambda(\mathrm{d}x, \mathrm{d}y) = \int_{X \otimes \mathbb{R}^d} f(x)\gamma(\mathrm{d}x, \mathrm{d}y).$$

It follows that, for every $f \in \mathfrak{F}$, $W \in \mathcal{W}$ and $k \geq 2$,

$$\int W(y_1, y_2, \dots, y_k)\lambda_f(\mathrm{d}y_1) \cdots \lambda_f(\mathrm{d}y_k) = \int W(y_1, y_2, \dots, y_k)\gamma_f(\mathrm{d}y_1) \cdots \gamma_f(\mathrm{d}y_k)$$

which implies from [12], Proof of Theorem 3.1, that for each $f \in \mathfrak{F}$, there is an $a = a(f) \in \mathbb{R}^d$ such that

$$\gamma_f(B) = \lambda_f(A + a). \tag{20}$$

We need to show that $a(f)$ is independent of f. We choose another $g \in \mathfrak{F}$ and let ϕ_f, ϕ_g and $\phi_{(f+g)/2} = \frac{1}{2}[\phi_f + \phi_g]$ denote the characteristic functions of λ_f, λ_g and $\frac{1}{2}[\lambda_f + \lambda_g]$, respectively. Likewise, ψ_f, ψ_g and $\psi_{(f+g)/2} = \frac{1}{2}[\psi_f + \psi_g]$ will denote the characteristic functions of γ_f, γ_g and $\frac{1}{2}[\gamma_f + \gamma_g]$, respectively. Then, (20) implies

$$\phi_f(t) = \psi_f(t)\mathrm{e}^{\mathbf{i}\langle t,a\rangle}; \qquad \phi_g(t) = \psi_g(t)\mathrm{e}^{\mathbf{i}\langle t,b\rangle}; \qquad \phi_{\frac{f+g}{2}}(t) = \psi_{\frac{f+g}{2}}(t)\mathrm{e}^{\mathbf{i}\langle t,c\rangle}.$$

Thus,

$$[\psi_f(t) + \psi_f(t)]\mathrm{e}^{i\langle t,c\rangle} = \phi_f(t) + \phi_g(t) = \psi_f(t)\mathrm{e}^{\mathbf{i}\langle t,a\rangle} + \psi_g(t)\mathrm{e}^{\mathbf{i}\langle t,b\rangle}$$

or, equivalently,

$$[\psi_f(t) + \psi_f(t)][\mathrm{e}^{\mathbf{i}\langle t,c\rangle} - 1] = \psi_f(t)[\mathrm{e}^{\mathbf{i}\langle t,a\rangle} - 1] + \psi_g(t)[\mathrm{e}^{\mathbf{i}\langle t,b\rangle} - 1].$$

Dividing both sides by t and letting $t \to 0$, we obtain $2c = a + b$, implying

$$\psi_f(t)[\mathrm{e}^{\mathbf{i}\langle t,c\rangle} - \mathrm{e}^{\mathbf{i}\langle t,a\rangle}] + \psi_g(t)[\mathrm{e}^{\mathbf{i}\langle t,c\rangle} - \mathrm{e}^{\mathbf{i}\langle t,b\rangle}] = 0$$

or

$$\psi_f(t)[\mathrm{e}^{\mathbf{i}\langle t,a-c\rangle} - 1] + \psi_g(t)[\mathrm{e}^{\mathbf{i}\langle t,b-c\rangle} - 1] = 0.$$

Since $a - c = c - b$,

$$\psi_g(t) = \psi_f(t)\frac{\mathrm{e}^{\mathbf{i}\langle t,b-c\rangle} - 1}{1 - \mathrm{e}^{-\mathbf{i}\langle t,a-c\rangle}} = \psi_f(t)\mathrm{e}^{\mathbf{i}\langle t,b-c\rangle}.$$

Starting from $f = 1$, it follows, that for every f, there is an $a(f)$ such that

$$\psi_f(t) = \psi_1(t)\mathrm{e}^{\mathbf{i}\langle t,a(f)\rangle}$$

and

$$\psi_1(t)\mathrm{e}^{i\langle t,a(\frac{1+f}{2})\rangle} = \psi_{\frac{1+f}{2}}(t) = \frac{1}{2}[\psi_1(t) + \psi_f(t))] = \frac{1}{2}[1 + \mathrm{e}^{\mathbf{i}\langle t,a(f)\rangle}]\psi_1(t).$$

Since $\psi_1(0) = 1$, we have for t near zero

$$\mathrm{e}^{\mathbf{i}\langle t,a(\frac{1+f}{2})\rangle} = \frac{1}{2}[1 + e^{\mathbf{i}\langle t,a(f)\rangle}],$$

which forces $a(f) = 0$ or $a = b = c$, proving the requisite claim. $\square$

2.3. *The compactness result.* The metric $\mathscr{D}$ and the resulting topology on $\mathscr{X}$ provides the following compactness result.

THEOREM 2.3. *Let $K \subset \mathscr{X}$ be such that as $\Theta = [\xi, \beta]$ varies over K, β ranges over a uniformly tight family (in the usual weak topology) in $\mathcal{M}_1(X)$. Then, every sequence from K has a subsequence that converges in the metric space $(\mathscr{X}, \mathscr{D})$.*

The proof of the above result needs the following notion of wide separation of measures. We remind the reader that for any measure $\gamma \in \mathcal{M}_{\leq 1}(X \otimes \mathbb{R}^d)$, $\gamma^{(2)}$ is the marginal of γ on $\mathbb{R}^d$. Then, two sequences of subprobability measures γ_n and δ_n on $X \otimes R^d$ are said to be *widely separated* if

$$\lim_{n\to\infty} \int V(y_1 - y_2)\gamma_n^{(2)}(\mathrm{d}y_1)\delta_n^{(2)}(\mathrm{d}y_2) = 0$$

for some strictly positive continuous V on $\mathbb{R}^d$ that tends to 0 at ∞. This requirement is equivalent to requiring (see [12], Lemma 2.4)

$$\lim_{n\to\infty} \int W(y_1, y_2)\gamma_n^{(2)}(\mathrm{d}y_1)\delta_n^{(2)}(\mathrm{d}y_2) = 0 \tag{21}$$

for all $W \in \mathfrak{C}_2$.

LEMMA 2.4. *Let $(\gamma_n)_n$ and $(\delta_n)_n$ be widely separated sequences in $\mathcal{M}_{\leq 1}(X \otimes \mathbb{R}^d)$. If f is bounded and continuous on X and $W \in \mathfrak{C}_k$ for some $k \geq 2$,*

$$\begin{aligned} \lim_{n\to\infty} \bigg| \int & f(x_1)\cdots f(x_k)W(y_1,\ldots,y_k)[\gamma_n+\delta_n](\mathrm{d}x_1,\mathrm{d}y_1)\cdots[\gamma_n+\delta_n](\mathrm{d}x_k,\mathrm{d}y_k) \\ &- \int f(x_1)\cdots f(x_k)W(y_1,\ldots,y_k)\gamma_n(\mathrm{d}x_1,\mathrm{d}y_1)\cdots\gamma_n(\mathrm{d}x_k,\mathrm{d}y_k) \\ &- \int f(x_1)\cdots f(x_k)W(y_1,\ldots,y_k)\delta_n(\mathrm{d}x_1,\mathrm{d}y_1)\cdots\delta_n(\mathrm{d}x_k,\mathrm{d}y_k)\bigg| = 0. \end{aligned} \tag{22}$$

PROOF. Let us expand the product $[\gamma_n + \delta_n](\mathrm{d}x_1, \mathrm{d}y_1)\cdots[\gamma_n + \delta_n](\mathrm{d}x_k, \mathrm{d}y_k)$. We need to show that the cross terms that involve any pair $\gamma_n(\mathrm{d}x_i, \mathrm{d}y_i)\cdots\delta_n(\mathrm{d}x_j, \mathrm{d}y_j)$ tend to 0 with n. For any k, any $W \in \mathfrak{C}_k$ and $1 \leq i < j \leq k$, there is a $V \in \mathfrak{C}_2$ such that $|W(y_1, \ldots, y_k)| \leq V(y_i - y_j)$. Then, from equation (21) it follows that

$$\begin{aligned} &\left| \int f(x_1)\cdots f(x_k)W(y_1,\ldots,y_k)\cdots\gamma_n(\mathrm{d}x_i,\mathrm{d}y_i)\cdots\delta_n(\mathrm{d}x_j,\mathrm{d}y_j)\cdots \right| \\ &\qquad\leq \int V(y_i - y_j)\gamma_n^{(2)}(\mathrm{d}y_i)\delta_n^{(2)}(\mathrm{d}y_j) \to 0. \end{aligned}$$

□

We now turn to the following proof.

PROOF OF THEOREM 2.3. The proof is carried out in several steps. We will choose subsequences repeatedly and use the diagonalization process. We will not use multilevel subscripts but will use n as the index in the sequences all the time. Let $\Theta_n = [\xi_n, \beta_n]$ be given so that $(\beta_n)_n$ is uniformly tight in the weak topology in $\mathcal{M}_1(X)$:

Step 1. By our assumption, choosing a subsequence we can assume that β_n converges weakly to β as probability distributions on X. Assume that ξ_n consists of a single orbit $\widetilde{\lambda}_n$. Let $q_n(\ell) = \sup_{a\in\mathbb{R}^d} \lambda_n^{(2)}[B(a,\ell)]$ where $\lambda_n \in \widetilde{\lambda}_n$ is any measure on the orbit and $B(a, \ell) = \{y \in \mathbb{R}^d : |y - a| \leq \ell\}$. Without loss of generality, by taking subsequences we can assume that $\lim_{n\to\infty} q_n(\ell) = q(\ell)$ exists for every positive integer ℓ. Let $q = \lim_{\ell\uparrow\infty} q(\ell)$. Assume that $p = \lim_{n\to\infty} \lambda_n(X \otimes \mathbb{R}^d)$ exists. Then, $q \le p$.

Step 2. If $q = 0$ and $\lambda_n^{(2)} \in \widetilde{\lambda}_n$, then $\lambda_n^{(2)} \hookrightarrow 0$ in the vague topology and

$$\lim_{n\to\infty} \int W(y_1, \dots, y_k)\lambda_n^{(2)}(dy_1) \cdots \lambda_n^{(2)}(dy_k) = 0$$

for every $W \in \mathcal{W}_k$. From the definition of $\mathscr{D}$, it is now easy to check that $\mathscr{D}(\Theta_n, \Theta) \to 0$ where $\Theta = [\varnothing, \beta]$.

Step 3. If $q = p$, then $\lambda_n \star \delta_{a_n} \Rightarrow \lambda$ weakly for a suitable choice of a_n. Again, we can verify that $\mathscr{D}(\Theta_n, \Theta) \to 0$ where $\Theta = [\xi, \beta]$ with ξ consisting of the single orbit $\widetilde{\lambda} = \{\lambda \star \delta_a\}$, $a \in \mathbb{R}^d$.

Step 4. Assume $0 < q < p$. Choose ℓ_0 so that $q(\ell_0) > \frac{3q}{4}$. Pick n_0 large enough such that for $n \ge n_0$, $q_n(\ell_0) > \frac{3q}{4}$. Since $q_n(\ell) = \sup_{\lambda\in\widetilde{\lambda}_n} \lambda(B(0, \ell))$, for $n \ge n_0$ we can find $\lambda_n \in \widetilde{\lambda}$ such that $\lambda_n^{(2)}[B(0, \ell_0)] \ge \frac{q}{2}$. Assume, by selecting a subsequence, that $\lambda_n^{(2)}$ has a vague limit α. Clearly, $\frac{q}{2} \le \alpha(\mathbb{R}^d) \le q$.

Since $\lim_{n\to\infty} \lambda_n^{(2)}(B(0, r)) = \alpha(B(0, r)) \le \alpha(\mathbb{R}^d)$, for any given r we can find n_r such that for $n \ge n_r$, $\lambda_n^{(2)}(B(0, r)) \le \alpha(\mathbb{R}^d) + \frac{1}{r}$. We can assume without loss of generality that n_r is increasing in r.

If $\rho_n = \{\sup r : n_r \le n\}$, then $\rho_n \to \infty$ and $\lambda_n^{(2)}[B(0, \rho_n)] \to \alpha(\mathbb{R}^d)$. The restriction of λ_n to $B(0, \rho_n)$ will converge weakly to α. Let us denote the restrictions of λ_n to $X \otimes B(0, \rho_n)$ and its complement $X \otimes [B(0, \rho_n)]^c$ by $\gamma_{1,n}$ and $\delta_{1,n}$, respectively, so that $\lambda_n = \gamma_{1,n} + \delta_{1,n}$. Then, a suitable subsequence of $\gamma_{1,n}$ will have a weak limit γ_1 on $X \otimes \mathbb{R}^d$ and $\lim_{n\to\infty} \int V(y_1 - y_2)\delta_{1,n}(\mathrm{d}y_1)\gamma_{1,n}(\mathrm{d}y_2) = 0$.

The orbit $\widetilde{\gamma}_1$ of γ_1 is a member of the set ξ. We work with $\delta_{1,n}$ and extract in the same way a limit $\widetilde{\gamma}_2$ and a leftover piece $\delta_{2,n}$. The process ends when $q = 0$ at some stage and we end up with a finite set ξ of $\{\widetilde{\gamma}_j\}$. If we do not end at a finite stage, we end up with a countable set. In either case, $\sum_{\widetilde{\gamma}_j\in\xi} \gamma_j^{(1)} \le \beta$.

Step 5. In estimating the distances $\mathbf{D}^*(\xi_1, \xi_2)$, if either ξ_1 or ξ_2 consists of several orbits $\{\widetilde{\lambda}\}$, we can afford to ignore orbits of measures with total mass at most ε. Since $k(r) \ge 2$, their total contribution to $\mathbf{D}^\star$ is at most

$$\sum_{r\ge2} \sum_{\substack{\widetilde{\lambda}: \\ \widetilde{\lambda}(X\otimes\mathbb{R}^d)\le\varepsilon}} \frac{[\widetilde{\lambda}(X \otimes \mathbb{R}^d)]^{k(r)}}{2^r} \le \sum_{\substack{\widetilde{\lambda}: \\ \widetilde{\lambda}(X\otimes\mathbb{R}^d)\le\varepsilon}} [\widetilde{\lambda}(X \otimes \mathbb{R}^d)]^2 \le \varepsilon \sum_{\substack{\widetilde{\lambda}: \\ \widetilde{\lambda}(X\otimes\mathbb{R}^d)\le\varepsilon}} \widetilde{\lambda}(X \otimes \mathbb{R}^d) \le \varepsilon.$$

We need to examine at most ε^{-1} orbits in each ξ_n. Taking a subsequence, we can assume that the number of such orbits m is the same in each ξ_n and link them as $\widetilde{\lambda}_{i,n}$ and deal with each sequence on its own. As limit along subsequences they will each generate a collection of orbits, and their union, denoted by ξ_ε, will again be a collection of orbits. In the end we can let $\varepsilon \to 0$ so that ξ_ε increases to a limiting collection of orbits ξ. Since we have uniform control over the combined contributions of all orbits with small masses, we can now pass to the limit $\varepsilon \to 0$. Lemma 2.4 will now allow us to show that we have convergence in the metric $\mathscr{D}$. From now on the details are identical to the proof of [12], Theorem 3.2, except we now have $X \otimes \mathbb{R}^d$ instead of $\mathbb{R}^d$. We omit the details to avoid repetition. □

The following result is an immediate consequence of Theorem 2.3.

COROLLARY 2.5. *Let $\{\Pi_n\}$ be a sequence of probability distributions on $\mathscr{X}$. For any $\Theta = [\xi, \beta]$ in $\mathscr{X}$, if the distributions $\{\widehat{\Pi}_n\}$ of β under Π_n is uniformly tight as a subset of $\mathcal{M}_1(X)$, then so is $\{\Pi_n\}$ as a subset of $\mathcal{M}_1(\mathscr{X})$.*

REMARK 1. From the definition (15) and (18) of the metric $\mathscr{D}$ on $\mathscr{X}$, it follows that for any continuous function V on $\mathbb{R}^d$ that tends to 0 at ∞, the function

$$\Psi(V,\Theta)=\sum_{\widetilde{\lambda}\in\xi}\int V(y_1-y_2)\widetilde{\lambda}^{(2)}(\mathrm{d}y_1)\widetilde{\lambda}^{(2)}(\mathrm{d}y_2)$$

is a bounded continuous function of $\Theta\in\mathscr{X}$, where $\Theta=[\xi,\beta]$, $\xi=\{\widetilde{\lambda}_j\}$ and $\widetilde{\lambda}^{(2)}$ is the marginal of any $\widetilde{\lambda}\in\xi$ on $\mathbb{R}^d$.

3. Stationary processes and processes with stationary increments.

3.1. *Some notation.* We start with the space $\Omega=\{\omega:(-\infty,\infty)\to\mathbb{R}^d:\omega(\cdot)\text{continuous}\}$ of $\mathbb{R}^d$-valued continuous functions on $(-\infty,\infty)$ which, equipped with the topology of uniform convergence on bounded intervals, is a complete separable metric space. The Borel σ-field of Ω, denoted by $\mathcal{F}$, is generated by $\{\omega(t):-\infty<t<\infty\}$. If $-\infty<a<b<\infty$, the σ-field $\mathcal{F}_{[a,b]}$ is the one generated by the increments $\{\omega(t)-\omega(s)\}$, $a\le s<t\le b$, and $\mathcal{F}_{\mathrm{inc}}=\sigma(\bigcup_{-\infty<a<b<\infty}\mathcal{F}_{[a,b]})$ is the σ-field of increments of ω. If we denote by

$$\Omega_0=\{\omega\in\Omega\in\Omega:\omega(0)=0\}\subset\Omega,$$

then Ω_0 can be identified (via a one-to-one map) with the *space of increments*, that is, continuous functions

$$h(s,t):\mathbb{R}\times\mathbb{R}\to\mathbb{R}^d\quad\text{such that } h(s,t)+h(t,u)=h(s,u).$$

Alternatively, we can also view the space of such functions $h(\cdot,\cdot)$ as equivalence classes $\Omega/\sim$ modulo constants, that is, for $\omega,\omega'\in\Omega$, we declare $\omega\sim\omega'$ if for some constant $a\in\mathbb{R}^d$, $\omega(t)=\omega'(t)+a$ for all $t\in\mathbb{R}$. Topology on Ω_0 is the natural one of uniform convergence on bounded subsets.

Next, we remark that we can identify Ω with $\Omega_0\otimes\mathbb{R}^d$ by mapping

$$\Omega\ni\omega\leftrightarrow(\omega',a)\quad\text{where } a=\omega(0),\omega'(t)=\omega(t)-\omega(0).$$

Thus, a subprobability measure $\mathbb{Q}$ on Ω can be viewed as a measure on $\Omega_0\times\mathbb{R}^d$ and will then have marginals $\mathbb{Q}^{(1)}\in\mathcal{M}_{\le 1}(\Omega_0)$, $\mathbb{Q}^{(2)}\in\mathcal{M}_{\le 1}(\mathbb{R}^d)$, respectively. Note that the marginal $\mathbb{Q}^{(1)}$ is just the distribution of the increments of a process that has $\mathbb{Q}$ for its distribution on $\mathcal{F}$.

3.2. *Stationary cocycles and group actions of $(S_t)_{t\in\mathbb{R}}$ on Ω and $(\theta_t)_{t\in\mathbb{R}}$ on $\Omega_0\otimes\mathbb{R}^d$.* We have the group of time translations $(S_t)_{t\in\mathbb{R}}$ acting on Ω as

$$S_t:\Omega\to\Omega,\qquad(S_t\omega)(s)=\omega(s+t),\tag{23}$$

while the group $(\theta_t)_{t\in\mathbb{R}}$ acts on $\Omega_0\times R^d$ as well as Ω_0 alone as

$$\begin{aligned}&\theta_t:\Omega_0\otimes\mathbb{R}^d\to\Omega_0\otimes\mathbb{R}^d\quad\text{with }(\theta_t)(\omega,x)=(\omega_t,x_t),\\&\text{where }\omega_t(s)=\omega(t+s)-\omega(t),\ \text{and } x_t=x+\omega(t)\quad\text{and}\\&\theta_t:\Omega_0\to\Omega_0\quad\text{with }(\theta_t\omega)(s)=\omega_t(s).\end{aligned}\tag{24}$$

Note that S_t-invariant probability measures on Ω are precisely θ_t-invariant probability measures on $\Omega_0\otimes\mathbb{R}^d$, are called *stationary processes* and are denoted by $\mathcal{M}_{\mathrm{s}}(\Omega_0\otimes\mathbb{R}^d)$. On the other hand, θ_t-invariant probability measures on Ω_0 are *processes with stationary increments* and are denoted by $\mathcal{M}_{\mathrm{si}}(\Omega_0)$. We can have a probability measure on $\Omega_0\otimes\mathbb{R}^d$ whose marginal on Ω_0 is θ_t-invariant, but it is not θ_t-invariant on $\Omega_0\times R^d$. These are precisely *nonstationary processes with stationary increments*. The following lemma provides a useful criterion that determines if a process β with stationary increments is the process of increments of a stationary process $\mathbb{Q}$. We will see that, even if it is, it is not unique, and, in fact, it is the *entire orbit* $\{\mathbb{Q}\star\delta_a:a\in\mathbb{R}^d\}$, where $\mathbb{Q}\star\delta_a$ is given by $(\mathbb{Q}\star\delta_a)[A\otimes B]=\mathbb{Q}[A\otimes(B-a)]$ on $\Omega_0\otimes\mathbb{R}^d$ and by $(\mathbb{Q}\star\delta_a)[A]=\mathbb{Q}[A-a]$ on Ω.

THEOREM 3.1. *Let β be an ergodic process with stationary increments (i.e., $\beta \in \mathcal{M}_{\mathrm{si}}(\Omega_0)$ is a θ_t-invariant and ergodic probability distribution on Ω_0). Then, either*

$$\lim_{\varepsilon\to 0} \mathbb{E}^\beta\left[\varepsilon \int_0^\infty \mathrm{e}^{-\varepsilon t} V\big(\omega(t)-\omega(0)\big)\,\mathrm{d}t\right] = 0 \tag{25}$$

for all continuous functions $V : \mathbb{R}^d \to \mathbb{R}$ with $\lim_{|x|\to\infty} |V(x)| = 0$ or there is a θ_t-invariant probability distribution $\mathbb{Q} \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$ such that $\beta = \mathbb{Q}^{(1)}$, that is, $\mathbb{Q}$ is a stationary process on Ω with β being the distribution of its increments.

REMARK 2. For the proof of the above result, it is enough to check the condition (25) for one strictly positive V.

PROOF OF THEOREM 3.1. Define $\mathbb{Q}_0(\mathrm{d}\omega\,\mathrm{d}y) = \beta(\mathrm{d}\omega) \otimes \delta_0(\mathrm{d}y) \in \mathcal{M}_1(\Omega_0 \otimes \mathbb{R}^d)$, and set

$$\nu_\varepsilon(\mathrm{d}\omega\,\mathrm{d}y) = \varepsilon \int_0^\infty \mathrm{d}t\mathrm{e}^{-\varepsilon t}\big(\mathbb{Q}_0\theta_t^{-1}\big)(\mathrm{d}\omega\,\mathrm{d}y). \tag{26}$$

Since the marginal $\nu_\varepsilon^{(1)}$ on Ω_0 of ν_ε is β for every $\varepsilon > 0$, the family $\{\nu_\varepsilon^{(1)}\}$ is (weakly) uniformly tight on Ω_0, and, by Lemma 2.1, $\{\nu_\varepsilon\}$ has wea-gue limit points. Let $\mathbb{Q}$ be any nonzero *weague* limit point of ν_ε as $\varepsilon \to 0$. We will show that, for any $\tau > 0$, $\theta_\tau\mathbb{Q} = \mathbb{Q}$. Then, $\mathbb{Q}$ on Ω will be stationary, and its marginal $\mathbb{Q}^{(1)}$ on Ω_0 will be dominated by β.

To show that $\theta_\tau\mathbb{Q} = \mathbb{Q}$, it is enough to verify that, for any $\tau > 0$ and any continuous $F : \Omega_0 \otimes \mathbb{R}^d \to \mathbb{R}$ satisfying

$$\lim_{|x|\to\infty} \sup_{\omega\in\Omega_0} \big|F(\omega, x)\big| = 0, \tag{27}$$

we have

$$\int_{\Omega_0}\int_{\mathbb{R}^d} F\big(\theta_\tau\omega, y+\omega(\tau)\big)\mathbb{Q}(\mathrm{d}\omega, \mathrm{d}y) = \int_{\Omega_0}\int_{\mathbb{R}^d} F(\omega, y)\mathbb{Q}(\mathrm{d}\omega\,\mathrm{d}y). \tag{28}$$

Actually, it is sufficient to prove the claim by taking F of the form $G(\omega)H(y)$ where G is bounded and continuous on Ω_0 and H is continuous and compactly supported on $\mathbb{R}^d$.

By construction, we note that, for any continuous and bounded function F on $\Omega_0 \otimes \mathbb{R}^d$,

$$\begin{aligned}
&\left|\int_0^\infty F\big(\theta_s(\cdot)\big)\varepsilon\mathrm{e}^{-\varepsilon s}\,\mathrm{d}s - \int_0^\infty F\big(\theta_{\tau+s}(\cdot)\big)\varepsilon\mathrm{e}^{-\varepsilon s}\,\mathrm{d}s\right| \\
&\quad = \left|\int_0^\tau F\big(\tau_s(\cdot)\big)\varepsilon\mathrm{e}^{-\varepsilon s}\,\mathrm{d}s - \int_t^\infty F\big(\tau_s(\cdot)\big)[\varepsilon\mathrm{e}^{-\varepsilon(s-\tau)} - \varepsilon\mathrm{e}^{-\varepsilon s}\,\mathrm{d}s\right| \\
&\quad \le \varepsilon\tau\mathrm{e}^{\varepsilon\tau}\|F\|_\infty + \|F\|_\infty\big[\mathrm{e}^{\varepsilon\tau} - 1\big],
\end{aligned}$$

which, together with the definition (26), imply that $\|\nu_\varepsilon - \theta_\tau\nu_\varepsilon\| \le |\mathrm{e}^{\tau\varepsilon} - 1| + \tau\varepsilon\mathrm{e}^{\tau\varepsilon}$. Since ν_ε converges wea-guely (along a subsequence) to $\mathbb{Q}$ to show that $\theta_\tau\mathbb{Q} = \mathbb{Q}$, it suffices to check that $\nu_\varepsilon\theta_\tau^{-1} \to \mathbb{Q}\theta_\tau^{-1}$ wea-guely, which is equivalent to showing

$$\int_{\Omega_0}\int_{\mathbb{R}^d} F\big(\theta_\tau\omega, y+\omega(\tau)\big)\nu_\varepsilon(\mathrm{d}\xi, \mathrm{d}y) \to \int_{\Omega_0}\int_{\mathbb{R}^d} F\big(\theta_\tau\omega, y+\omega(\tau)\big)\mathbb{Q}(\mathrm{d}\omega, \mathrm{d}y) \tag{29}$$

for any continuous $F : \Omega_0 \times \mathbb{R}^d \to \mathbb{R}$ satisfying (27). The distribution of $\omega(\tau)$ under ν_ε is that of the original increment $\omega(\tau) - \omega(0)$ under β, and, so given $\delta > 0$, there is a function $g(z)$ with compact support in $\mathbb{R}^d$ with $0 \le g \le 1$ such that, for all $\varepsilon > 0$,

$$\int \big(1 - g\big(\omega(\tau)\big)\big)\nu_\varepsilon^{(1)}(\mathrm{d}\omega) = \int \big(1 - g\big(\omega(\tau)\big)\big)\beta(\mathrm{d}\omega) \le \delta \tag{30}$$

and since $\mathbb{Q}^{(1)} \le \beta$, we have $\int (1 - g(\omega(\tau)))\mathbb{Q}^{(1)}(\mathrm{d}\omega) \le \delta$. If we replace $F(\theta_\tau \omega, y + \omega(\tau))$ by $G(\omega, a) = g(\omega(\tau))F(\theta_\tau\omega, y + \omega(\tau))$, since $\omega(\tau)$ is now restricted to a bounded set, we have

$$\lim_{|y|\to\infty} \sup_{\omega} |G(\omega, y)| = 0$$

and (29) follows with the help of (28) and (30). Thus, $\theta_\tau \mathbb{Q} = \mathbb{Q}$ for all $\tau > 0$.

We will now conclude. Note that the measures ν_ε all have marginals $\nu_\varepsilon^{(1)} = \beta$, the wea-gue limit $\mathbb{Q}$ of ν_ε is θ_t-invariant and $\mathbb{Q}^{(1)}$ is dominated by β. Since β is ergodic, it is equal to $c\beta$ for some $0 \le c \le 1$, and $\frac{1}{c}\mathbb{Q}$ is a stationary process with $\beta = \frac{1}{c}\mathbb{Q}^{(1)}$, provided we can show that $c > 0$. However, if $c = 0$, we have

$$\lim_{\varepsilon\to 0} \int_{\Omega_0}\int_{\mathbb{R}^d} F(\omega, y)\nu_\varepsilon(\mathrm{d}\omega, \mathrm{d}y) = 0$$

for continuous functions F satisfying (27). In particular, $F(\omega, y)$ can be any continuous function $V(y)$ satisfying $\lim_{|y|\to\infty} |V(y)| = 0$, and, hence,

$$\begin{aligned}\int_{\Omega_0}\int_{\mathbb{R}^d} V(y)\nu_\varepsilon(\mathrm{d}\omega, \mathrm{d}y) &= \int_{\mathbb{R}^d} V(y)\nu_\varepsilon^{(2)}(\mathrm{d}y)\\ &= \mathbb{E}^\beta\left[\varepsilon\int_0^\infty \mathrm{e}^{-\varepsilon t} V\big(\omega(t) - \omega(0)\big)\,\mathrm{d}t\right] \to 0\end{aligned}$$

as $\varepsilon \to 0$. □

In order to apply Theorem 3.1 in the present context, we need another fact, whose proof can be found in [3], Theorem 4.1.

LEMMA 3.2. *Let $\mathcal{Y}$ be a Polish space and $(S_t)_{t\ge 0}$ a one-parameter family of homeomorphisms of $\mathcal{Y}$. For each* $\mathrm{y} \in \mathcal{Y}$, *let*

$$\nu_\varepsilon(\mathrm{y}, \cdot) = \varepsilon \int_0^\infty \mathrm{e}^{-\varepsilon t}\delta_{S_t(\mathrm{y})}\,\mathrm{d}t,$$

that is, for any $\mathrm{y} \in \mathcal{Y}$ *and* $A \subset \mathcal{Y}$, $\nu_\varepsilon(\mathrm{y}, A) = \int_0^\infty \mathrm{d}t\,\varepsilon\mathrm{e}^{-\varepsilon t}\mathbb{1}\{S_t(\mathrm{y}) \in A\}$. *Then, the following implications hold*:

- *For each* $\varepsilon > 0$, $\mathrm{y} \to \nu_\varepsilon(\mathrm{y}, \cdot)$ *is a map from* $X \to \mathcal{M}_1(\mathcal{Y})$, *and any weak limit point* $\nu = \lim_{\varepsilon\to 0} \nu_\varepsilon$ *is invariant under* S_t.
- *Let* μ_n *be a sequence of* S_t*-invariant probability measures on* $\mathcal{Y}$ *converging weakly to* μ, *and* $\lambda_{\varepsilon,\mu_n}$ *be the distribution of* $\nu_\varepsilon(\mathrm{y}, \cdot)$ *on* $\mathcal{M}_1(\mathcal{Y})$, *that is,*

  $$\lambda_{\varepsilon,\mu_n}(G) = \mu_n\big[\mathrm{y} \in \mathcal{Y}\colon \nu_\varepsilon(y, \cdot) \in G\big] \quad \forall G \subset \mathcal{M}_1(\mathcal{Y}).$$

  Then, $(\lambda_{\varepsilon,\mu_n})_n$ *is uniformly tight on* $\mathcal{M}_1(\mathcal{Y})$, *and, for any sequence* $\varepsilon_n \to 0$, *any weak limit point* λ *of* $\lambda_{\varepsilon_n,\mu_n}$ *is supported on the space* $\mathcal{M}_{\mathrm{inv}}(\mathcal{Y})$ *of* S_t*-invariant probability measures on* $\mathcal{Y}$. *Moreover,*

  $$\int_{\mathcal{M}_{\mathrm{inv}}(\mathcal{Y})} \nu\lambda(\mathrm{d}\nu) = \mu.$$

The following corollary is an immediate consequence of Lemma 3.2 for the particular case $\mathcal{Y} = \Omega_0 \otimes \mathbb{R}^d$.

COROLLARY 3.3. *Let $\mathbb{Q}_n$ be a sequence of invariant probability measures on $\Omega_0 \otimes \mathbb{R}^d$ such that their marginals $\mathbb{Q}_n^{(1)} \in \mathcal{M}_1(\Omega_0)$ converge weakly to a limit $\beta \in \mathcal{M}_1(\Omega_0)$. For any $\omega \in \Omega_0$ and $\varepsilon > 0$, let*

$$\nu_\varepsilon(\omega, \cdot) = \nu_\varepsilon\big((\omega, 0), \cdot\big) = \varepsilon \int_0^\infty \mathrm{e}^{-\varepsilon t} \delta_{\theta_t(\omega,0)}\, \mathrm{d}t \in \mathcal{M}_1\big(\Omega_0 \otimes \mathbb{R}^d\big). \tag{31}$$

Let ε_n be any sequence such that $\varepsilon_n \to 0$, and $\Pi_n \in \mathcal{M}_1(\mathcal{M}_1(\Omega_0 \times \mathbb{R}^d))$ is the distribution of ν_{ε_n} under $\mathbb{Q}_n$, that is, $(\omega, 0) \in \Omega_0 \otimes \mathbb{R}^d$ is sampled according to $\mathbb{Q}_n$; for any $\mathscr{A} \subset \mathcal{M}_1(\Omega_0 \otimes \mathbb{R}^d)$, we set

$$\Pi_n(\mathscr{A}) \stackrel{\text{(def)}}{=} \mathbb{Q}_n\big[\nu_{\varepsilon_n}\big((\omega, 0), \cdot\big) \in \mathscr{A}\big]. \tag{32}$$

Then, the projection $\widehat{\Pi}_n \in \mathcal{M}_1(\mathcal{M}_1(\Omega_0))$ of Π_n is uniformly tight, any weak limit point $\widehat{\Pi} = \lim_{n\to\infty} \widehat{\Pi}_n$ is supported on the space of θ_t-invariant (but not necessarily ergodic) distributions $\mathcal{M}_{\mathrm{si}}(\Omega_0)$, and, moreover,

$$\lim_{n\to\infty} \mathbb{Q}_n^{(1)} = \beta = \int_{\mathcal{M}_{\mathrm{si}}} \nu \widehat{\Pi}(\mathrm{d}\nu).$$

REMARK 3. Note that the space $\mathcal{M}_{\mathrm{si}}(\Omega_0)$, the space of processes with stationary increments, consists of two sets $\mathcal{M}_{\mathrm{coc}}$ and $\mathcal{M}_{\mathrm{ncoc}}$ corresponding to those that are increments of a stationary process and those that are not. Thus, Corollary 3.3 implies that the limit β (in Corollary 3.3) can be written as the sum $\beta = \beta_{\mathrm{coc}} + \beta_{\mathrm{ncoc}}$, where

$$\beta_{\mathrm{coc}} = \int_{\mathcal{M}_{\mathrm{coc}}} \nu \Pi(\mathrm{d}\nu) \quad \text{and} \quad \beta_{\mathrm{ncoc}} = \int_{\mathcal{M}_{\mathrm{ncoc}}} \nu \Pi(\mathrm{d}\nu).$$

For any $\nu \in \mathcal{M}_{\mathrm{coc}}$, we have a stationary distribution $\widetilde{\alpha}_\nu = \{\alpha_\nu \star \delta_a : a \in \mathbb{R}^d\}$ that are the possible marginals.

4. Relative entropy estimates and variational arguments.

4.1. *Relative entropy of processes with stationary increments.* From now on we shall assume that $d = 3$. Let us recall that, for any $a < b$, we denote by $\mathcal{F}_{[a,b]}$ the σ-algebra generated by all increments $\{\omega(s) - \omega(r) : a \le r < s \le b\}$, and $\mathbb{P}$ refers to the law of three-dimensional Brownian increments. For any process $\mathbb{Q} \in \mathcal{M}_{\mathrm{si}}(\Omega_0)$ with stationary increments, if $\mathbb{Q}_{0,\omega}$ denotes the regular conditional probability distribution of $\mathbb{Q}$ given $\mathcal{F}_{(-\infty,0)}$, which is the σ-field generated by $\{(\omega(t) - \omega(s)) : -\infty < s < t \le 0\}$, then

$$H_t(\mathbb{Q}|\mathbb{P}) = \mathbb{E}^{\mathbb{Q}}\big[h_{\mathcal{F}_{[0,t]}}(\mathbb{Q}_{0,\omega}|\mathbb{P})\big] \tag{33}$$

defines the entropy of the increments of the process $\mathbb{Q}$ with respect to $\mathbb{P}$ over the σ-field $\mathcal{F}_{[0,t]}$. Here, for any two probability measures μ and ν on any σ-algebra of the form $\mathcal{F} = \mathcal{F}_{[a,b]}$ on Ω, we denote by

$$h_{\mathcal{F}}(\mu|\nu) = \sup_f \left\{ \int f\, \mathrm{d}\mu - \log\left(\int \mathrm{e}^f\, \mathrm{d}\nu \right) \right\} \tag{34}$$

the relative entropy of the probability measure μ with respect to ν on $\mathcal{F}$, and the supremum above is taken over all continuous, bounded and $\mathcal{F}$-measurable functions. For our purposes it is useful to collect some well-known properties of $H_t(\mathbb{Q}|\mathbb{P})$.

LEMMA 4.1. *Either $H_t(\mathbb{Q}|\mathbb{P}) \equiv \infty$ for all $t > 0$, or, $H_t(\mathbb{Q}|\mathbb{P}) = t H(\mathbb{Q}|\mathbb{P})$, where the map $H(\cdot|\mathbb{P}) : \mathcal{M}_{\mathrm{si}} \to [0, \infty]$, called the (specific) relative entropy of a process $\mathbb{Q}$ with stationary increments, satisfies the following properties*:

- *$H(\cdot|\mathbb{P})$ is convex and lower semicontinuous in the usual weak topology.*
- *$H(\cdot|\mathbb{P})$ is coercive (i.e., for any $\ell \geq 0$, the sublevel sets $\{\mathbb{Q}\colon H(\mathbb{Q}|\mathbb{P}) \leq \ell\}$ are weakly compact as measures on Ω_0.*
- *The map $\mathbb{Q} \mapsto H(\mathbb{Q}|\mathbb{P})$ is, in fact, linear. In particular, for any probability measure Γ on $\mathcal{M}_{\mathrm{si}}$,*

$$H\biggl(\int \mathbb{Q}\Gamma(\mathrm{d}\mathbb{Q})|\mathbb{P}\biggr) = \int H(\mathbb{Q}|\mathbb{P})\Gamma(\mathrm{d}\mathbb{Q}). \tag{35}$$

PROOF. We refer to [2], Section 3, for the proofs of the above assertions. Indeed, the fact that either $H_t(\mathbb{Q}|\mathbb{P}) \equiv \infty$ for all $t > 0$ or, $H_t(\mathbb{Q}|\mathbb{P}) = tH(\mathbb{Q}|\mathbb{P})$ can be found in [2], Theorem 3.1. The convexity and lower semicontinuity of the $H(\cdot|\mathbb{P})$ is proved in [2], Theorem 3.3, while coercivity of this map is a result of the variational representation of $H_t(\cdot|\mathbb{P})$ proved in [2], Theorem 3.2. Finally, the linearity of $H(\cdot|\mathbb{P})$ is shown in [2], Theorem 3.5. □

Finally, we remark that, if $\mathbb{Q} \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$ is a stationary process, we can also consider the σ-fields $\Sigma_{[a,b]}$ generated by $\{\omega(t) : a \leq t \leq b\}$ and define

$$\widehat{H}_t(\mathbb{Q}|\mathbb{P}) = \mathbb{E}^{\mathbb{Q}}\bigl[h_{\Sigma_{[0,t]}}(\mathbb{Q}_{0,\omega}|\mathbb{P})\bigr],$$

where $\mathbb{Q}_{0,\omega}$ is now the conditional probability given $\Sigma_{(-\infty,0]}$. Since a stationary process is also one with stationary increments, both $H_t(\mathbb{Q}|\mathbb{P})$ (defined in (33)) and $\widehat{H}_t(\mathbb{Q}|\mathbb{P})$ make sense for $\mathbb{Q} \in \mathcal{M}_s$, and the same statements as in Lemma 4.1 continue to hold in this case too; in fact, it is not difficult to see that these two objects coincide, that is,

LEMMA 4.2. *For any $\mathbb{Q} \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \otimes \mathbb{R}^d)$,*

$$H(\mathbb{Q}|\mathbb{P}) = \widehat{H}(\mathbb{Q}|\mathbb{P}). \tag{36}$$

PROOF. Indeed, since by (35) in Lemma 4.1 both H and $\widehat{H}$ are linear, we can assume that $\mathbb{Q}$ is stationary and ergodic (else $\mathbb{Q}$ can be written as a mixture of stationary ergodic measures). If either $H(\mathbb{Q})$ or $\widehat{H}(\mathbb{Q})$ is finite, then (33)–(34) imply that $\mathbb{E}^{\mathbb{Q}}[|\omega(1) - \omega(0)|] < \infty$. By the ergodic theorem, we have

$$\omega(0) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \bigl(\omega(0) - \omega(t)\bigr)\,dt + E^{\mathbb{Q}}[\omega(0)],$$

and, since there is no essential difference between the two σ-fields $\mathcal{F}$ and Σ, we have the desired equality (36). □

4.2. *Relative entropy estimates*. The following lemma allows us to apply the results from Section 2 and Section 3 to the singular function $x \mapsto 1/|x|$ in $\mathbb{R}^3$.

LEMMA 4.3. *For any $\eta > 0$, let*

$$\begin{aligned} &V(x) = \frac{1}{|x|}, \qquad V_\eta(x) = \frac{1}{(\eta^2 + |x|^2)^{1/2}} \quad \textit{and} \\ &Y_\eta(x) = (V - V_\eta)(x). \end{aligned} \tag{37}$$

Then,

$$\lim_{\eta\to 0} \sup_{\substack{\mathbb{Q}\in\mathcal{M}_{\mathrm{si}} \\ H(\mathbb{Q}|\mathbb{P})\leq C}} \sup_{0<\varepsilon\leq 1} \varepsilon\mathbb{E}^{\mathbb{Q}}\biggl[\int_0^\infty \mathrm{d}t\mathrm{e}^{-\varepsilon t} Y_\eta\bigl(\omega(t) - \omega(0)\bigr)\biggr] = 0, \tag{38}$$

where in the supremum above $C<\infty$ is any finite constant. Moreover, given any $\delta>0$, there is a constant $C(\delta)$ such that for all $\varepsilon\le 1$ and $\mathbb{Q}\in\mathcal{M}_{\mathrm{si}}(\Omega_0)$ with $H(\mathbb{Q}|\mathbb{P})<\infty$,

$$\mathbb{E}^{\mathbb{Q}}\biggl[\varepsilon\int_0^\infty \frac{\mathrm{e}^{-\varepsilon t}}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\biggr]\le \delta H(\mathbb{Q}|\mathbb{P})+C(\delta). \tag{39}$$

PROOF. With any $\rho>0$, we estimate

$$\begin{aligned}
F(t)&\stackrel{\text{(def)}}{=}\mathbb{E}^{\mathbb{Q}}\biggl[\int_0^t Y_\eta\bigl(\omega(t)-\omega(0)\bigr)\,\mathrm{d}t\biggr]\\
&\le \frac{t}{\rho}H(\mathbb{Q}|\mathbb{P})+\frac{1}{\rho}\log\mathbb{E}^{\mathbb{P}}\biggl[\exp\biggl(\rho\int_0^t Y_\eta\bigl(\omega(s)-\omega(0)\bigr)\,\mathrm{d}s\biggr)\biggr]\\
&\le \frac{t}{\rho}H(\mathbb{Q}|\mathbb{P})+\frac{1}{\rho}\log\mathbb{E}^{\mathbb{P}}\biggl[\exp\biggl(\rho\int_0^\infty Y_\eta\bigl(\omega(s)-\omega(0)\bigr)\,\mathrm{d}s\biggr)\biggr]\\
&\le \frac{t}{\rho}H(\mathbb{Q}|\mathbb{P})+\frac{1}{\rho}\log\biggl[\frac{1}{1-\rho\sup_{x\in\mathbb{R}^3}\mathbb{E}^{\mathbb{P}}[\int_0^\infty Y_\eta(x+\omega(s)-\omega(0))\,\mathrm{d}s]}\biggr]\\
&=\frac{t}{\rho}H(\mathbb{Q}|\mathbb{P})-\frac{1}{\rho}\log\bigl[1-\rho h(\eta)\bigr].
\end{aligned}\tag{40}$$

In the first upper bound in the above display, we have used the relative entropy inequality,

$$\mathbb{E}^{\mathbb{Q}}\biggl[\int_0^t V\bigl(\omega(s)-\omega(0)\bigr)\,\mathrm{d}s\biggr]\le\frac{tH(\mathbb{Q}|\mathbb{P})}{\rho}+\frac{1}{\rho}\log\mathbb{E}^{\mathbb{P}}\biggl[\exp\biggl(\rho\int_0^t V\bigl(\omega(s)-\omega(0)\bigr)\,\mathrm{d}s\biggr)\biggr],$$

and in the third upper bound we have used Khasminski's lemma stating that, for $V\ge 0$,

$$\mathbb{E}^{\mathbb{P}}\biggl[\exp\biggl(\rho\int_0^\infty V\bigl(\omega(s)-\omega(0)\bigr)\,\mathrm{d}s\biggr)\biggr]\le\frac{1}{1-\rho\sup_{x\in\mathbb{R}^3}\mathbb{E}^{\mathbb{P}}[\int_0^\infty V(x+\omega(s)-\omega(0))\,\mathrm{d}s]}.$$

It is not hard to check that the supremum in

$$\sup_{x\in\mathbb{R}^3}\mathbb{E}^{\mathbb{P}}\biggl[\int_0^\infty\frac{1}{|x+\omega(s)-\omega(0)|^{\frac32}}\,ds\biggr]$$

is attained at $x=0$. Furthermore, with (37) we can estimate $0\le Y_\eta(x)\le\frac{c\sqrt{\eta}}{|x|^{\frac32}}$, and, therefore,

$$\begin{aligned}
h(\eta)&=E^{\mathbb{P}}\biggl[\int_0^\infty Y_\eta\bigl(\omega(s)-\omega(0)\bigr)\,ds\biggr]\le c\sqrt{\eta}\int_{R^3}\int_0^\infty\frac{1}{|x|^{\frac32}}\frac{1}{(2\pi t)^{\frac32}}e^{-\frac{|x|^2}{2t}}\,\mathrm{d}t\,\mathrm{d}x\\
&=c'\sqrt{\eta}\to 0
\end{aligned}$$

as $\eta\to 0$. For $\varepsilon\le 1$ and $H(\mathbb{Q}|\mathbb{P})\le C$, the expectation on the left-hand side in (38) can be estimated with the help of (40) and repeated integration by parts, which yields

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\biggl[\varepsilon\int_0^\infty \mathrm{d}t\mathrm{e}^{-\varepsilon t}Y_\eta\bigl(\omega(t)-\omega(0)\bigr)\biggr]&=\varepsilon\int_0^\infty\mathrm{e}^{-\varepsilon t}F'(t)\,\mathrm{d}t\\
&=\varepsilon^2\int_0^\infty e^{-\varepsilon t}F(t)\,\mathrm{d}t\\
&\le\frac{1}{\rho}H(\mathbb{Q}|\mathbb{P})-\frac{\varepsilon}{\rho}\log\bigl[1-\rho h(\eta)\bigr]\\
&\le\frac{C}{\rho}-\frac{1}{\rho}\log\bigl[1-\rho h(\eta)\bigr].
\end{aligned}$$

We can now let $\eta \to 0$ followed by $\rho \to \infty$ to deduce (38). To prove (39), given any $\delta > 0$ in the previous step, we can take η to be small enough so that $h(\eta) < \frac{\delta}{2}$. Then, with $\rho = \delta^{-1}$ and $V_\eta(x) \le \frac{1}{\eta}$, we have

$$\mathbb{E}^{\mathbb{Q}}\biggl[\varepsilon \int_0^\infty \frac{\mathrm{e}^{-\varepsilon t}}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\biggr] \le \delta H(\mathbb{Q}|\mathbb{P}) + \frac{1}{\eta} + \delta \log 2 = \delta H(\mathbb{Q}|\mathbb{P}) + C(\delta)$$

which proves (39). $\square$

Combining the above result with Theorem 3.1, we now have:

COROLLARY 4.4. *Let $\mathbb{Q}$ be an ergodic process with stationary increments such that $H(\mathbb{Q}|\mathbb{P}) < \infty$. Then, either*

$$\lim_{\varepsilon\to 0} \mathbb{E}^{\mathbb{Q}}\biggl[\varepsilon \int_0^\infty \frac{\mathrm{e}^{-\varepsilon t}}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\biggr] = 0 \tag{41}$$

or there is a stationary process on Ω with $\mathbb{Q}$ being the distribution of its increments.

LEMMA 4.5. *If $\mathbb{Q} \in \mathcal{M}_{\mathrm{s}}(\Omega_0 \times \mathbb{R}^3)$ is a stationary process with marginal $\mu \in \mathcal{M}_1(\mathbb{R}^3)$ and $H(\mathbb{Q}|\mathbb{P}) < \infty$, then there exists a nonnegative function $\psi \in H^1(\mathbb{R}^3)$ with $\|\psi\|_2 = 1$ such that $\mathrm{d}\mu = [\psi(x)]^2\,\mathrm{d}x$. Moreover, if ψ is strictly positive and $\log\psi$ is twice continuously differentiable with bounded derivatives, we also have*

$$H(\mathbb{Q}|\mathbb{P}) = \frac{1}{2}\int_{\mathbb{R}^3} \bigl|(\nabla\psi)(x)\bigr|^2\,\mathrm{d}x + H(\mathbb{Q}|\widetilde{\mathbb{Q}}_\psi), \tag{42}$$

where $\widetilde{\mathbb{Q}}_\psi \in \mathcal{M}_1(\Omega_0 \otimes \mathbb{R}^3)$, the stationary Markov process with generator $\frac{1}{2}\Delta + \frac{\nabla\psi}{\psi}\cdot\nabla$ and marginal $\psi^2(x)\,\mathrm{d}x \in \mathcal{M}_1(\mathbb{R}^3)$.

REMARK 4. The notation $\widetilde{\mathbb{Q}}_\psi$ for the process with generator $\frac{1}{2}\Delta + \frac{\nabla\psi}{\psi}\cdot\nabla$ will be justified later; see Theorem 4.9 and Remark 6 below.

PROOF OF LEMMA 4.5. Note that the first statement of the lemma is immediate. Assume that ψ is strictly positive and $\log\psi$ has the required regularity. Let $\widetilde{\mathbb{Q}}_\psi^\omega$ and $\mathbb{Q}^\omega$ be, respectively, the conditional distributions of $\widetilde{\mathbb{Q}}_\psi$ and $\mathbb{Q}$ on $\mathcal{F}_{[0,1]}$ given $\mathcal{F}_{(-\infty,0)}$. If we write

$$\frac{\mathrm{d}\mathbb{Q}^\omega}{\mathrm{d}\mathbb{P}} = \biggl(\frac{\mathrm{d}\mathbb{Q}^\omega}{\mathrm{d}\widetilde{\mathbb{Q}}_\psi^\omega}\biggr)\biggl(\frac{\mathrm{d}\widetilde{\mathbb{Q}}_\psi^\omega}{\mathrm{d}\mathbb{P}}\biggr),$$

we need to check that

$$\mathbb{E}^{\mathbb{Q}}\biggl[\log\biggl(\frac{\mathrm{d}\widetilde{\mathbb{Q}}_\psi^\omega}{\mathrm{d}\mathbb{P}}\biggr)\biggr] = \frac{1}{2}\int_{R^3}\bigl|(\nabla\psi)(x)\bigr|^2\,\mathrm{d}x. \tag{43}$$

Applying Girsanov's formula followed by Itô's formula, we have

$$\log\biggl(\frac{\mathrm{d}\widetilde{\mathbb{Q}}_\psi}{\mathrm{d}\mathbb{P}}\biggr) = \int_0^1 \biggl(\frac{\nabla\psi}{\psi}\biggr)(\omega(t))\cdot\mathrm{d}\omega(t) - \frac{1}{2}\int_0^1\biggl|\frac{\nabla\psi}{\psi}\biggr|^2(\omega(t))\,\mathrm{d}t$$

$$= \log\psi(\omega(1)) - \log\psi(\omega(0)) \tag{44}$$

$$- \frac{1}{2}\int_0^1 (\Delta\log\psi)(\omega(t))\,\mathrm{d}t - \frac{1}{2}\int_0^1 \biggl|\frac{\nabla\psi}{\psi}\biggr|^2(\omega(t))\,\mathrm{d}t. \tag{45}$$

Then (43) follows by taking expectation on both sides in the last display with respect to the stationary process with marginal density $\psi^2(x)\,\mathrm{d}x$ and integration by parts. This expectation of the right hand side above then reduces to

$$-\frac{1}{2}\int_{\mathbb{R}^3}(\Delta\log\psi)(x)\psi^2(x)\,\mathrm{d}x-\frac{1}{2}\int_{\mathbb{R}^3}\frac{|\nabla\psi|^2(x)}{\psi^2(x)}\psi^2(x)\,\mathrm{d}x$$
$$=\frac{1}{2}\int_{\mathbb{R}^3}|\nabla\psi|^2(x)\,\mathrm{d}x. \qquad \square$$

REMARK 5. In the context of proving Theorem 1.1 for the Fröhlich polaron, we will use Lemma 4.5, when ψ is the maximizer ψ_0 of the Pekar variational problem (6) and, in this case, $\log\psi_0$ has the required regularity according to [8]. Indeed, let us check the assumption on the maximizer ψ_0. Recall that

$$g_0=\sup_{\substack{\psi\in H^1(\mathbb{R}^3)\\ \|\psi\|_2=1}}\left\{\int\int_{\mathbb{R}^3\times\mathbb{R}^3}\frac{\psi^2(x)\psi^2(y)}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y-\frac{1}{2}\int_{\mathbb{R}^3}\mathrm{d}x\big|\nabla\psi(x)\big|^2\right\}.$$

A simple perturbation argument shows that the maximizing function $\psi_0\in H^1(\mathbb{R}^3)$ satisfies the Euler–Lagrange equation

$$\left(\Delta+4\int_{\mathbb{R}^3}\frac{\psi_0^2(y)}{|x-y|}\,\mathrm{d}y\right)\psi_0(x)=\lambda\psi_0(x). \tag{46}$$

We multiply (46) on both sides by $\psi_0(x)$, integrate over $\mathbb{R}^3$ and recall that $\int_{\mathbb{R}^3}\psi_0^2=1$, to see that

$$\lambda=4\int\int_{\mathbb{R}^3\times\mathbb{R}^3}\frac{\psi_0^2(x)\psi_0^2(y)}{|x-y|}-\|\nabla\psi_0\|_2^2$$
$$=2\left(2\int\int_{\mathbb{R}^3\times\mathbb{R}^3}\frac{\psi_0^2(x)\psi_0^2(y)}{|x-y|}-\frac{1}{2}\|\nabla\psi_0\|_2^2\right)\geq 2g_0>0.$$

It has been shown by Lieb [8], Theorem 8, (iii), that, if $\lambda>0$ in (46), then the maximizing function $\psi_0\in C^\infty$ goes to zero at infinity and, hence, ψ_0 is a strong solution of (46).

4.3. *Coercivity of the variational formulas and tightness of maximizers.*

LEMMA 4.6. *Consider the variational problem*

$$g(\varepsilon)=\sup_{\mathbb{Q}\in\mathcal{M}_{\mathrm{si}(\Omega_0)}}\left[\mathbb{E}^{\mathbb{Q}}\left(\varepsilon\int_0^\infty e^{-\varepsilon t}\frac{1}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\right)-H(\mathbb{Q}|\mathbb{P})\right]. \tag{47}$$

Then, the supremum is attained over a nonempty set $\mathfrak{m}_\varepsilon$ of processes with stationary increments.

The proof of the above result is based on the following well-known fact which crucially exploits that the independence of the increments of the underlying measure $\mathbb{P}$.

LEMMA 4.7. *Fix any $C<\infty$. Then, the set $\{\mathbb{Q}\in\mathcal{M}_{\mathrm{si}}(\Omega_0)\colon H(\mathbb{Q}|\mathbb{P})\leq C\}$ is uniformly tight in the weak topology.*

PROOF OF LEMMA 4.6. Let us consider any maximizing sequence $(\mathbb{Q}_n)_n \subset \mathcal{M}_{\mathrm{si}}(\Omega_0)$ such that

$$\left[E^{\mathbb{Q}_n}\left[\varepsilon \int_0^\infty \mathrm{e}^{-\varepsilon t} \frac{1}{|\omega(t)-\omega(0)|}\,dt\right] - H(\mathbb{Q}_n|\mathbb{P})\right] \to g(\varepsilon).$$

By (39), for any $\delta \in (0,1)$ there is a $C = C(\delta)$ such that, for all n,

$$\mathbb{E}^{\mathbb{Q}_n}\left[\varepsilon \int_0^\infty \mathrm{e}^{-\varepsilon t} \frac{1}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\right] \le \delta H(\mathbb{Q}_n|\mathbb{P}) + C(\delta)$$

for all $\varepsilon \le 1$. Combining the last two estimates, imply that, for some $C < \infty$, $\sup_n H(\mathbb{Q}_n|\mathbb{P}) \le C$ and then by Lemma 4.7, $\{\mathbb{Q}_n\}_n$ is also uniformly tight in $\mathcal{M}_{\mathrm{si}}(\Omega_0)$. If

$$\mathbb{Q} \stackrel{(\mathrm{def})}{=} \lim_{n\to\infty} \mathbb{Q}_n$$

is any subsequential limit point, the lower semicontinuity of $H(\cdot|\mathbb{P})$ then implies that $\mathbb{Q} \in \mathfrak{m}_\varepsilon$. □

LEMMA 4.8. *For any $\varepsilon > 0$, let $\mathfrak{m}_\varepsilon$ be the maximizers of the variational formula* (47). *Then, for any $\varepsilon_0 > 0$, $\bigcup_{0<\varepsilon<\varepsilon_0} \mathfrak{m}_\varepsilon$ is uniformly tight in $\mathcal{M}_{\mathrm{si}}(\Omega_0)$.*

PROOF. Since $g(\varepsilon) \ge 0$ for all $\varepsilon > 0$, if $\mathbb{Q} \in \bigcup_{\varepsilon>0} \mathbf{m}_\varepsilon$, then again (39), for any $\delta \in (0,1)$ implies

$$0 \le E^{\mathbb{Q}}\left[\varepsilon \int_0^\infty \frac{\mathrm{e}^{-\varepsilon t}}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\right] - H(\mathbb{Q}|\mathbb{P}) \le (\delta - 1)H(\mathbb{Q}|\mathbb{P}) + C(\delta),$$

providing a uniform upper bound

$$H(\mathbb{Q}|\mathbb{P}) \le \frac{C(\delta)}{1-\delta}$$

on $H(\mathbb{Q}|P)$ which together with Lemma 4.7 proves the lemma. □

4.4. *A key result*: *Identification of any limiting maximizer.* We will now prove a key result which identifies any limiting maximizer of the variational problem $\mathfrak{m}_\varepsilon$ as the increments of the Pekar process. Combined with Lemma 4.8 it constitutes the main argument for the convergence of the polaron measure $\widehat{\mathbb{P}}_\varepsilon$ to the increments of the Pekar process.

THEOREM 4.9. *Let $\varepsilon_n \to 0$ and $\mathbb{Q}_n \subset \mathbf{m}_{\varepsilon_n}$ so that $\mathbb{Q}_n \Rightarrow \mathbb{Q}$ on $\mathcal{M}_1(\Omega_0)$. Then, $\mathbb{Q}$ is the distribution of the increments of a stationary process $\widetilde{\mathbb{Q}}_\psi \in \mathcal{M}_1(\Omega_0 \otimes \mathbb{R}^3)$, which is the diffusion corresponding to the generator*

$$\frac{1}{2}\Delta + \frac{\nabla\psi}{\psi}\cdot\nabla,$$

with invariant density $\psi^2(x)\,dx$, where ψ is any maximizer of

$$\sup_{\|\psi\|_2=1}\left[\int_{R^3}\int_{\mathbb{R}^3} \frac{\psi^2(x)\psi^2(y)}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\int_{\mathbb{R}^3} |(\nabla\psi)(x)|^2\,\mathrm{d}x\right].$$

REMARK 6. Recall that the maximizer ψ for the above variational problem is unique only up to translations. This is reflected in the fact that $\widetilde{\mathbb{Q}}_\psi = \{\mathbb{Q}_{\psi_0} \star \delta_a : a \in \mathbb{R}^3 \in \widetilde{\mathcal{M}}_1(\Omega_0 \otimes \mathbb{R}^3)$ is an orbit under translations, and *any representative* $\mathbb{Q}_{\psi_0} \star \delta_a$ of that orbit has $\mathbb{Q} := \lim_{n\to\infty} \mathbb{Q}_n \in \mathcal{M}_1(\Omega_0)$ as the distribution of its increments.

PROOF OF THEOREM 4.9. The proof is carried out in several steps.

Step 1. It is known that ([3])

$$\lim_{\varepsilon\to 0} g(\varepsilon) = g_0 = \sup_{\|\psi\|_2=1}\left[\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\psi^2(x)\psi^2(y)}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\int_{\mathbb{R}^3} \left|(\nabla\psi)(x)\right|^2\mathrm{d}x\right], \tag{48}$$

and, therefore, for $(\mathbb{Q}_n) \subset \mathfrak{m}_{\varepsilon_n}$, we have

$$\left[\mathbb{E}^{\mathbb{Q}_n}\left[\varepsilon_n \int_0^\infty e^{-\varepsilon_n t} \frac{1}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\right] - H(\mathbb{Q}_n|\mathbb{P})\right] \to g_0.$$

Step 2. Note that, $x \mapsto 1/|x|$ is an even function in $\mathbb{R}^3$. Moreover, since each $\mathbb{Q}_n$ is a process with stationary increments (i.e., θ_t invariant measure on Ω_0), for each n and $\varepsilon_n > 0$, we have the identity

$$\int_0^\infty \mathrm{d}s\,\varepsilon_n \mathrm{e}^{-\varepsilon_n s}(\mathbb{Q}_n\theta_s^{-1})(\cdot) = \int_0^\infty \mathrm{d}s\,\varepsilon_n \mathrm{e}^{-\varepsilon_n s}\mathbb{Q}_n(\cdot) = \mathbb{Q}_n(\cdot).$$

Using these two facts, we have

$$\begin{aligned} &\mathbb{E}^{\mathbb{Q}_n}\left[\varepsilon_n \int_0^\infty \frac{\mathrm{e}^{-\varepsilon_n t}}{|\omega(t)-\omega(0)|}\,\mathrm{d}t\right] \\ &\quad= \mathbb{E}^{\mathbb{Q}_n}\left[\frac{\varepsilon_n}{2}\int_{-\infty}^\infty \frac{\mathrm{e}^{-\varepsilon_n|t|}}{|\omega(t)-\omega(0)|}\,dt\right] \\ &\quad= \mathbb{E}^{\mathbb{Q}_n}\left[\varepsilon_n^2 \int_0^\infty\int_0^\infty \frac{\mathrm{e}^{-\varepsilon_n t-\varepsilon_n s}}{|\omega(t)-\omega(s)|}\,\mathrm{d}t\,\mathrm{d}s\right] \\ &\quad= \mathbb{E}^{\Pi_n}\left[\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{1}{|x_1-x_2|}\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)\right], \end{aligned} \tag{49}$$

where in the last display we used the definition for the averages $\nu_{\varepsilon_n}(\cdot,\cdot)$ from (31) and, as in (32), we wrote $\Pi_n \in \mathcal{M}_1(\mathcal{M}_1(\Omega_0 \otimes \mathbb{R}^3))$ for the distribution of ν_{ε_n} under $\mathbb{Q}_n$. Then, $\lambda_n^{(2)} \in \mathcal{M}_1(\mathbb{R}^3)$ is distributed according to the marginal of Π_n on $\mathcal{M}_1(\mathbb{R}^3)$, and, moreover, the double integral

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{1}{|x_1-x_2|}\lambda^{(2)}(\mathrm{d}x_1)\lambda^{(2)}(\mathrm{d}x_2)$$

is a function of the orbit $\widetilde{\lambda}^{(2)} \in \widetilde{\mathcal{M}}_1(\mathbb{R}^3)$, justifying the last identity in (49).

Now, according to Corollary 3.3, the projection $\widehat{\Pi}_n$ of Π_n on $\mathcal{M}_1(\Omega_0)$ is a uniformly tight family, and, therefore, by Theorem 2.3 and Corollary 2.5, Π_n itself is uniformly tight on $\mathcal{M}_1(\mathscr{X})$. Moreover, by Remark 1, for any continuous function $V : \mathbb{R}^3 \to \mathbb{R}$ vanishing at infinity the function

$$\Psi(V,[\xi,\beta]) = \sum_{\widetilde{\lambda}\in\xi}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} V(x_1-x_2)\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)$$

is continuous on the metric space $(\mathscr{X},\mathscr{D})$. Furthermore, the same argument as in the proof of Lemma 4.6 provides a uniform bound in $H(\mathbb{Q}_n|\mathbb{P})$ allowing us to invoke the uniform estimate (38) to control the unboundedness of the Coulomb potential. This estimate, combined with the above continuity of $[\xi,\beta] \mapsto \Psi(V_\eta,[\xi,\beta])$ (with $V_\eta(x) = (|x|^2+\eta^2)^{-1/2}$), shows that if we take the limit

$$\Pi = \lim_n \Pi_n$$

in the weak topology on $\mathcal{M}_1(\mathscr{X})$ along a subsequence, then (49) dictates

$$\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in S}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)\Big]-H(\mathbb{Q}|\mathbb{P})\geq g_0. \tag{50}$$

In the above assertion we have additionally used the lower semicontinuity of $H(\cdot|\mathbb{P})$.

Step 3. Recall that typical elements of $\mathscr{X}$ are denoted by $[\xi,\beta]$, where $\beta\in\mathcal{M}_1(\Omega_0)$ and $\xi=\{\widetilde{\lambda}_j\}_j$, any $\lambda\in\xi$ is an orbit of a measure with total mass at most 1 on $\Omega_0\otimes\mathbb{R}^3$ and $\widetilde{\lambda}^{(1)}$ and $\widetilde{\lambda}^{(2)}$ are their projections on Ω_0 and $\mathbb{R}^3$, respectively. Let $\Pi\in\mathcal{M}_1(\mathscr{X})$ denote the limit of $\Pi_n=\mathbb{Q}_n\nu_{\varepsilon_n}^{-1}$ from Step 2 above. Then, for any $[\xi,\beta]\in\mathscr{X}$, which is distributed according to Π, we have

$$\sum_{\widetilde{\lambda}\in\xi}\widetilde{\lambda}^{(1)}(\cdot)\leq\beta(\cdot)\quad \Pi\text{-a.s.},$$

and we denote the difference by $\beta_0(\cdot)=\beta(\cdot)-\sum_{\widetilde{\lambda}\in\xi}\widetilde{\lambda}^{(1)}(\cdot)$ which is again θ_t-invariant on Ω_0. We can write β as a convex combination of probability measures on Ω_0,

$$\beta(\cdot)=m(\beta_0)\overline{\beta}_0(\cdot)+\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda}^{(1)})\overline{\lambda}^{(1)}(\cdot)\quad\text{where}$$

$$\overline{\lambda}^{(1)}(\cdot)=\frac{1}{m(\widetilde{\lambda}^{(1)})}\widetilde{\lambda}^{(1)}(\cdot)\quad\text{and}\quad\overline{\beta}_0(\cdot)=\frac{1}{m(\beta_0)}\beta_0(\cdot),$$

where $m(\widetilde{\lambda}^{(1)})$ and $m(\beta_0)$ are the total masses of $\widetilde{\lambda}^{(1)}$ and β_0, respectively. This convex decomposition, combined with the linearity of the map $H(\cdot|\mathbb{P})$, we have

$$H(\mathbb{Q}|\mathbb{P})=\mathbb{E}^{\Pi}\Big[m(\beta_0)H\big(\overline{\beta}_0\big|\mathbb{P}\big)+\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})H\big(\overline{\lambda}^{(1)}\big|\mathbb{P}\big)\Big]\geq\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})H\big(\overline{\lambda}^{(1)}|\mathbb{P}\big)\Big]. \tag{51}$$

On the other hand, since $m(\widetilde{\lambda})\leq 1$, we have

$$\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)\Big] \tag{52}$$

$$\leq\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\overline{\lambda}^{(2)}(\mathrm{d}x_1)\overline{\lambda}^{(2)}(\mathrm{d}x_2)\Big]. \tag{53}$$

Step 4. Putting (50), (51) and (52) together and observing that $\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})\leq 1$, we have:

$$g_0\leq\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)\Big]-H(\mathbb{Q}|\mathbb{P}) \tag{54}$$

$$\leq\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}[m(\widetilde{\lambda})]^2\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\overline{\lambda}^{(2)}(\mathrm{d}x_1)\overline{\lambda}^{(2)}(\mathrm{d}x_2)-\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})H\big(\overline{\lambda}^{(1)}|\mathbb{P}\big)\Big] \tag{55}$$

$$\leq\mathbb{E}^{\Pi}\Big[\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\overline{\lambda}^{(2)}(\mathrm{d}x_1)\overline{\lambda}^{(2)}(\mathrm{d}x_2)-\sum_{\widetilde{\lambda}\in\xi}m(\widetilde{\lambda})H\big(\overline{\lambda}^{(1)}|\mathbb{P}\big)\Big] \tag{56}$$

$$\leq\sup_{\widetilde{\lambda}\in\widetilde{\mathcal{M}}_s(\Omega_0\otimes\mathbb{R}^3)}\Big[\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x_1-x_2|}\widetilde{\lambda}^{(2)}(\mathrm{d}x_1)\widetilde{\lambda}^{(2)}(\mathrm{d}x_2)-H\big(\widetilde{\lambda}^{(1)}|\mathbb{P}\big)\Big] \tag{57}$$

$$\leq\sup_{\substack{\psi\in H^1(\mathbb{R}^3)\\ \|\psi\|_2=1}}\Big[\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{1}{|x-y|}\psi^2(x)\psi^2(y)\,\mathrm{d}y-\frac{1}{2}\int_{\mathbb{R}^3}\big|\nabla\psi(x)\big|^2\,\mathrm{d}x\Big] \tag{58}$$

$$=g_0$$

In the penultimate step we have used the fact from (42) that if $\widetilde{\mathbb{Q}}$ is a stationary process with marginal distribution $\psi^2(x)\,\mathrm{d}x$, then

$$H(\widetilde{\mathbb{Q}}^{(1)}|\mathbb{P}) = \frac{1}{2}\int_{R^3}\big|(\nabla\psi)(x)\big|^2\,dx + H(\widetilde{\mathbb{Q}}^{(1)}|\mathbb{Q}_\psi^{(1)}) \geq \frac{1}{2}\int_{\mathbb{R}^3}\big|(\nabla\psi)(x)\big|^2\,dx \tag{59}$$

and equality above holds only when $\widetilde{\mathbb{Q}}^{(1)} = \mathbb{Q}_\psi$ which is the common distribution of increments of the stationary diffusion with generator $\frac{1}{2}\Delta + \frac{\nabla\psi_y}{\psi_y}\cdot\nabla$.

Step 5. Note that equality should hold at every step between (54) and (58). Then, it is easy to see that Π is concentrated on a single orbit $[\{\widetilde{\lambda}\},\lambda]\in\mathscr{X}$ and λ is the diffusion $\mathbb{Q}_\psi$, corresponding to a ψ that maximizes (48). Indeed, recall Corollary 4.4. Since equality in (56) forces $m(\widetilde{\lambda})$ to be 0 or 1, but, since the sum over $\widetilde{\lambda}$ is at most 1, there is only one orbit $\widetilde{\lambda}$ in ξ; by the equality in (58) and previous remark regarding equality in (59), the stationary process $\lambda\in\mathcal{M}_{\mathrm{s}}(\Omega_0\otimes\mathbb{R}^d)$ must be the diffusion $\mathbb{Q}_\psi$ with generator $\frac{1}{2}\Delta + \frac{\nabla\psi_y}{\psi_y}\cdot\nabla$, with ψ being a maximizer of (48). This argument concludes the proof of Theorem 4.9. □

5. Identification of the strong coupling polaron and Kac interactions: Proofs of Theorem 1.1 and Theorem 1.2.

5.1. *Proof of Theorem* 1.1. Recall the definition of the Polaron measure $\widehat{\mathbb{P}}_{\varepsilon,T}$ from (3). In [13], Theorem 5.1, it was shown that, for all sufficiently small (or sufficiently large) $\varepsilon>0$, the limit $\widehat{\mathbb{P}}_\varepsilon = \lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}\in\mathcal{M}_{\mathrm{si}}(\Omega_0)$ exists. The following result provides an explicit identification of $\widehat{\mathbb{P}}_\varepsilon$ in the strong coupling limit $\lim_{\varepsilon\to0}\widehat{\mathbb{P}}_\varepsilon$ which, combined with [13], Theorem 5.1, will also complete the proof of Theorem 1.1.

THEOREM 5.1. $\lim_{\varepsilon\to0}\widehat{\mathbb{P}}_\varepsilon = \mathbb{Q}_\psi$ *exists and is equal to the distribution of increments of the stationary diffusion with generator*

$$\frac{1}{2}\Delta + \frac{\nabla\psi_y}{\psi_y}\cdot\nabla.$$

As already remarked earlier, the distribution of increments of a stationary diffusion with the above generator is independent of $y\in\mathbb{R}^3$. Combining Theorem 4.9 with Lemma 4.8 will conclude the proof of the above result, once we show that, for $\varepsilon>0$ (small enough, but fixed), the infinite volume polaron measure $\widehat{\mathbb{P}}_\varepsilon\in\mathfrak{m}_\varepsilon$.

PROPOSITION 5.2. *Let* $\widehat{\mathbb{P}}_\varepsilon = \lim_{T\to\infty}\widehat{\mathbb{P}}_{\varepsilon,T}$ *and* $\mathfrak{m}_\varepsilon$ *be the set of maximizers of the variational problem* (47). *Then*, $\widehat{\mathbb{P}}_\varepsilon\in\mathfrak{m}_\varepsilon$.

The proof of the above result depends on a *strong large deviation principle* for the distribution of the empirical process

$$R_T(\omega,\cdot) = \frac{1}{T}\int_0^T \delta_{\theta_s\omega_T(\cdot)}\,\mathrm{d}s\in\mathcal{M}_{\mathrm{si}}(\Omega_0) \tag{60}$$

of increments, where

$$\omega_T(s) = \begin{cases}\omega(s) & \text{if } 0\leq s\leq T,\\ n\omega(T)+\omega(r) & \text{if } s=nT+r, n\in\mathbb{N}, 0\leq r<T.\end{cases}$$

Let $\mathbb{Q}_T = \mathbb{P}R_T^{-1}\in\mathcal{M}_1(\mathcal{M}_{\mathrm{si}}(\Omega_0))$ be the distribution of the empirical process under three-dimensional Brownian increments $\mathbb{P}$. Then:

LEMMA 5.3. *The family $(\mathbb{Q}_T)_{T>0}$ satisfies a strong large deviation principle as $T \to \infty$ in the space of probability measures on $\mathcal{M}_{\mathrm{si}}(\Omega_0)$ with rate function $H(\cdot|\mathbb{P})$. In other words,*

$$\liminf_{T\to\infty} \frac{1}{T} \log \mathbb{Q}_T(G) \geq -\inf_{\mathbb{Q}\in G} H(\mathbb{Q}|\mathbb{P}) \quad \forall G \subset \mathcal{M}_{\mathrm{si}} \text{ open},$$
$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{Q}_T(F) \leq -\inf_{\mathbb{Q}\in F} H(\mathbb{Q}|\mathbb{P}) \quad \forall F \subset \mathcal{M}_{\mathrm{si}} \text{ closed}. \tag{61}$$

PROOF. The proof of the lower bound for all open sets $G \subset \mathcal{M}_{\mathrm{si}}$ and the upper bound for all compact sets $K \subset \mathcal{M}_{\mathrm{si}}$ in (61) follows directly from the arguments of [2] modulo slight changes, and the details are omitted. To strengthen the upper bound to all closed sets $C \subset \mathcal{M}_{\mathrm{si}}$, it suffices to show exponential tightness for the distributions $\mathbb{Q}_T$ that requires, for any $\ell > 0$, existence of a compact set $K_\ell \subset \mathcal{M}_{\mathrm{si}}$ so that

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{Q}_T[K_\ell^{\mathrm{c}}] \leq -\ell. \tag{62}$$

To prove the above claim, for $i = 1, 2, 3$, if we set

$$\|\omega_i^\star\| = \sup_{0\leq s\leq t\leq 1} \frac{|\omega_i(s) - \omega_i(t)|}{|s-t|^{1/4}},$$

then by Borell's inequality, for some constants $C_1, C_2 > 0$,

$$\mathbb{P}[\|\omega_i^\star\| > \lambda] \leq C_1 \exp\left[-\frac{\lambda^2}{2C_2}\right]$$

and, consequently, $\mathbb{E}^{\mathbb{P}}[\mathrm{e}^{\|\omega_i^\star\|}] < \infty$. Then,

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{E}^{\mathbb{Q}_T}\left[\mathrm{e}^{T\sum_{i=1}^3 \|\omega_i^\star\|}\right] < \infty,$$

and the desired exponential tightness (62) follows readily, proving the requisite upper bound in (61) for all closed sets. □

The following lemma will complete the proof of Proposition 5.2.

LEMMA 5.4. *Fix any $\varepsilon > 0$. Then, the distributions $\widehat{\mathbb{P}}_{\varepsilon,T} R_T^{-1}$ of the empirical process of increments under the polaron measure satisfies a strong large deviation principle in $\mathcal{M}_1(\mathcal{M}_{\mathrm{si}}(\Omega_0))$ with rate function*

$$\mathbb{Q} \mapsto g(\varepsilon) - \left[\mathbb{E}^{\mathbb{Q}}\left(\varepsilon \int_0^\infty \frac{\varepsilon \mathrm{e}^{-\varepsilon r}\,\mathrm{d}r}{|\omega(r) - \omega(0)|}\right) - H(\mathbb{Q}|\mathbb{P})\right].$$

Moreover, any limit point of $\widehat{\mathbb{P}}_{\varepsilon,T} R_T^{-1}$ (as $T \to \infty$) lies in $\mathfrak{m}_\varepsilon$.

PROOF. It follows from Lemma 5.3 that $\lim_{T\to\infty} \frac{1}{T} \log Z_{\varepsilon,T} = g(\varepsilon)$ (we can again invoke Lemma 4.3 to control the unboundedness of the Coulomb potential in $\mathbb{R}^3$) and the required strong large deviation principle for $\widehat{\mathbb{P}}_{\varepsilon,T} R_T^{-1}$ follows immediately. In particular, the upper bound of this large deviation implies that, for any open neighborhood $U(\mathfrak{m}_\varepsilon)$ of $\mathfrak{m}_\varepsilon$, $\limsup_{T\to\infty} \frac{1}{T} \log \widehat{\mathbb{P}}_{\varepsilon,T}[R_T \notin U(\mathfrak{m}_\varepsilon)] < 0$ which, together with compactness of the level sets of $H(\cdot|\mathbb{P})$, forces all limit points of $\widehat{\mathbb{P}}_{\varepsilon,T} R_T^{-1}$ to be supported in $\mathfrak{m}_\varepsilon$. □

PROOF OF PROPOSITION 5.2 AND THEOREM 5.1. Since the actual limit $\widehat{\mathbb{P}}_\varepsilon = \lim_{T\to\infty} \widehat{\mathbb{P}}_{\varepsilon,T}$ exists ([13], Theorem 5.1), it follows, together with Lemma 5.4, that $\widehat{\mathbb{P}}_\varepsilon \in \mathfrak{m}_\varepsilon$. Theorem 4.9 then concludes the proof of Theorem 5.1. □

5.2. *Proof of Theorem* 1.2. Recall the definition of the path measure $\widehat{\mathbb{P}}^{(V)}_{\varepsilon,T}$ from (1) defined w.r.t. any $V:\mathbb{R}^d\to\mathbb{R}$ which is rotationally symmetric, continuous and vanishes at infinity. For such a function V in Theorem 1.2, we also assume that the variational problem

$$\sup_{\substack{\psi\in H^1(\mathbb{R}^d)\\ \|\psi\|_2=1}}\left[\int\int_{\mathbb{R}^{2d}} V(x-y)\psi^2(x)\psi^2(y)\,\mathrm{d}x\,\mathrm{d}y-\frac{1}{2}\|\nabla\psi\|_2^2\right]$$

admits a smooth maximizer $\psi^{(V)}$ which is unique, modulo translations in $\mathbb{R}^d$. For the proof of Theorem 1.2, we can essentially repeat the argument for proving Theorem 1.1, except that we can no longer appeal to [13], Theorem 5.1, to justify the existence of the thermodynamic limit

$$\lim_{T\to\infty}\widehat{\mathbb{P}}^{(V)}_{\varepsilon,T}=:\widehat{\mathbb{P}}^{(V)}_{\varepsilon}. \tag{63}$$

The reason for this is that the method in our earlier article [13] hinges upon the Gaussian representation $\frac{1}{|x|}=\sqrt{\frac{2}{\pi}}\int_0^\infty \mathrm{e}^{-\frac{u^2|x|^2}{2}}\,\mathrm{d}u$ of the particular choice of the Coulomb potential $V(x)=\frac{1}{|x|}$ in $d=3$. This representation is no longer available for a general interaction V being considered in Theorem 1.2. However, in [11] we considered Gibbs measures with translation invariance of the form

$$\mathrm{d}\widehat{\mathbb{P}}^{(H)}_{\alpha,T}:=\frac{1}{Z_T}\exp\left\{\alpha\int_{-T}^{T}\int_{-T}^{T}H\big(t-s,\omega(t)-\omega(s)\big)\,\mathrm{d}s\,\mathrm{d}t\right\}\mathrm{d}\mathbb{P},$$

defined w.r.t. the law $\mathbb{P}$ of increments of d-dimensional Brownian paths. In the main result of [11], it has been shown that, for any coupling parameter $\alpha>0$, the thermodynamic limit $\lim_{T\to\infty}\widehat{\mathbb{P}}^{(H)}_{\alpha,T}$ exists and a central limit theorem also holds for $\lim_{T\to\infty}\widehat{\mathbb{P}}^{(H)}_{\alpha,T}[\frac{\omega(T)-\omega(-T)}{\sqrt{2T}}\in\cdot]$, provided that the interaction H satisfies

$$\sup_{x\in\mathbb{R}^d}\big|H(t,x)\big|\le C\big(1+|t|\big)^{-\gamma}\quad\text{for some }\gamma>2,\ \text{and }C\in(0,\infty). \tag{64}$$

In the context of Theorem 1.2, we have

$$H(t,x)=\varepsilon\mathrm{e}^{-\varepsilon|t|}V(x)\quad\text{with }\|V\|_\infty\le C\in(0,\infty).$$

Clearly, the above choice satisfies the desired requirement (64) for any fixed $\varepsilon>0$, and we can appeal to [11] to justify (63). Given the existence of the limit $\widehat{\mathbb{P}}^{(V)}_{\varepsilon}$ in (63) and repeating the large deviation arguments leading to Proposition 5.2 (with $V(\cdot)$ replacing the Coulomb potential $1/|\cdot|$), we have that $\widehat{\mathbb{P}}^{(V)}_{\varepsilon}$ is a maximizer of the variational problem (47), again with $V(\omega(t)-\omega(0))$ replacing the Coulomb potential $\frac{1}{|\omega(t)-\omega(0)|}$ therein. Now, we can exactly repeat the arguments for the proof of Theorem 1.1 to conclude the proof of Theorem 1.2 as well. Certainly, owing to the boundedness of V, some arguments now simplify (e.g., the truncation arguments needed to handle the singularity of the Coulomb potential carried out in Lemma 4.3 will no longer be required for a bounded function V).

Acknowledgments. The authors are very grateful to Herbert Spohn for helpful discussions and comments on a first draft of the manuscript. The first author was supported in part by Deutsche Forschungsgemeinschaft (DFG) under Germany's Excellence Strategy EXC 2044–390685587, Mathematics Münster: Dynamics–Geometry–Structure.

REFERENCES

[1] BOLTHAUSEN, E., KÖNIG, W. and MUKHERJEE, C. (2017). Mean-field interaction of Brownian occupation measures II: A rigorous construction of the Pekar process. *Comm. Pure Appl. Math.* **70** 1598–1629. MR3666565 https://doi.org/10.1002/cpa.21682

[2] DONSKER, M. D. and VARADHAN, S. R. S. (1983). Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.* **36** 183–212. MR0690656 https://doi.org/10.1002/cpa.3160360204

[3] DONSKER, M. D. and VARADHAN, S. R. S. (1983). Asymptotics for the polaron. *Comm. Pure Appl. Math.* **36** 505–528. MR0709647 https://doi.org/10.1002/cpa.3160360408

[4] DYBALSKI, W. and SPOHN, H. (2020). Effective mass of the Polaron-revisited. *Ann. Henri Poincaré* **21** 1573–1594. MR4087365 https://doi.org/10.1007/s00023-020-00892-7

[5] FEYNMAN, R. P. (1972). *Statistical Mechanics*. Benjamin, Reading. MR1634071

[6] KÖNIG, W. and MUKHERJEE, C. (2017). Mean-field interaction of Brownian occupation measures, I: Uniform tube property of the Coulomb functional. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 2214–2228. MR3729652 https://doi.org/10.1214/16-AIHP788

[7] LANDAU, L. D. (1933). Electron motion in crystal lattices. *Phys. Z. Sowet.* **3** 664–665.

[8] LIEB, E. H. (1976/77). Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation. *Stud. Appl. Math.* **57** 93–105. MR0471785 https://doi.org/10.1002/sapm197757293

[9] LIEB, E. H. and SEIRINGER, R. (2020). Divergence of the effective mass of a polaron in the strong coupling limit. *J. Stat. Phys.* **180** 23-33. MR4130981 https://doi.org/10.1007/s10955-019-02322-3

[10] LIEB, E. H. and THOMAS, L. E. (1997). Exact ground state energy of the strong-coupling polaron. *Comm. Math. Phys.* **183** 511–519. MR1462224 https://doi.org/10.1007/s002200050040

[11] MUKHERJEE, C. (2017). Central limit theorem for Gibbs measures including long range and singular interactions and homogenization of the stochastic heat equation. Preprint, arXiv:1706.09345.

[12] MUKHERJEE, C. and VARADHAN, S. R. S. (2016). Brownian occupation measures, compactness and large deviations. *Ann. Probab.* **44** 3934–3964. MR3572328 https://doi.org/10.1214/15-AOP1065

[13] MUKHERJEE, C. and VARADHAN, S. R. S. (2020). Identification of the polaron measure I: Fixed coupling regime and the central limit theorem for large times. *Comm. Pure Appl. Math.* **73** 350–383. MR4054359 https://doi.org/10.1002/cpa.21858

[14] PEKAR, S. I. (1949). Theory of polarons. *Ž. èksp. Teor. Fiz.* **19**.

[15] SPOHN, H. (1987). Effective mass of the polaron: A functional integral approach. *Ann. Physics* **175** 278–318. MR0887976 https://doi.org/10.1016/0003-4916(87)90211-9