

CONTENTS

1. The setting	1
2. Motivation and statement of the main result	1
3. Proof of the main theorem	2
3.1. A random walk representation	3
3.2. Reducing the problem to a Markov chain	4
3.3. Ventsel's LDP for paths	5
3.4. An application of the contraction principle	7
3.5. Evaluating the rate function	8
4. Implications of the result	10
4.1. Equivalence of SSEP and QHF	10
4.2. SSEP $\leftrightarrow$ QHF correspondences	11
References	12

1. THE SETTING

Consider N bosons on the complete graph Λ_V on V vertices, interacting via hard-core repulsion. More precisely we consider the Hamiltonian of a general quantum lattice gas on Λ_V

$$H_{V,N}\Psi(x_1, \dots, x_N) := - \sum_{i=1}^N \sum_{y \in \Lambda_V} c(x_i, y) [\Psi(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_N) - \Psi(x_1, \dots, x_N)] \quad (1.1) \quad \{\text{eq:line1}\}$$

$$+ U(x_1, \dots, x_N) \Psi(x_1, \dots, x_N), \quad (1.2) \quad \{\text{eq:line2}\}$$

acting on $\mathcal{H}_{V,N} := \{\phi(x) | x \in \Lambda_V\}^{\otimes N}$, where we set

$$c(x, y) \equiv \frac{1}{V} \quad \text{and} \quad U(x_1, \dots, x_N) := \begin{cases} 0 & \text{if } i \neq j \Rightarrow x_i \neq x_j \\ \infty & \text{if otherwise.} \end{cases} \quad (1.3)$$

Remark 1.1 (Interpretation of the Hamiltonian). *The first line (1.1) corresponds to the N -fold tensor product of the normalised discrete Laplacian Δ_{Λ_V} on the complete graph, i.e. $-\Delta_{\Lambda_V} \phi(x) = \frac{1}{V} \sum_{y \in \Lambda_V} (\phi(x) - \phi(y))$. Hence this line describes N non-interacting particles, each following an independent simple random walk on the complete graph.*

The second line (1.2) encodes the hard-core repulsion constraint; with this we can think of the trajectories of the particles as random walks conditioned on avoiding each other. This will be made more precise later.

The main result of [1] is the calculation of the *canonical free energy per unit volume*

$$\beta f(\rho, \beta) := - \lim_{\substack{N, V \rightarrow \infty \\ N/V \rightarrow \rho}} \frac{1}{V} \log Z(\beta, N, V), \quad (1.4) \quad \{\text{eq:thermoli}\}$$

where Z is the canonical partition function and defined as usual as

$$Z(\beta, N, V) := \text{Tr}_{\text{sym}} e^{-\beta H_{V,N}}.$$

Here Tr_{sym} is the trace over the symmetric subspace of $\mathcal{H}_{V,N}$. (This essentially encodes the fact that we are considering *bosons*.)

2. MOTIVATION AND STATEMENT OF THE MAIN RESULT

As pointed out in Remark 1.1, setting $U \equiv 0$ would give independent simple random walks which are considerably easier to analyse. In fact it can be shown that in this case (i.e. considering a non-interacting mean-field system of bosons) the free energy $f_{0,B}$ is given by

$$\beta f_{0,B}(\rho, \beta) = \begin{cases} -(\rho + 1) \log(\rho + 1) + \rho \log \rho + \beta \rho & \text{if } \rho < \rho_{0,c} = \frac{1}{e^\beta - 1}, \\ -(\rho_{0,c} + 1) \log(\rho_{0,c} + 1) + \rho_{0,c} \log \rho_{0,c} + \beta \rho_{0,c} & \text{if } \rho \geq \rho_{0,c} \end{cases}, \quad (2.1)$$

while the free energy $f_{0,F}$ for the same model, but with fermions instead of bosons, is given by

$$\beta f_{0,F}(\rho, \beta) = \rho \log \rho + (1 - \rho) \log(1 - \rho) + \beta \rho. \quad (2.2)$$

Note how *on the complete graph* the *free* Bose gas exhibits a phase transition (this is commonly referred to as the Bose-Einstein condensation (BEC)), whereas the corresponding Fermi system does not. The situation is qualitatively different on graphs that are not as connected as the complete graph: Since the one-dimensional hard-core Bose system can be mapped onto a one-dimensional non-interacting fermion gas there is no BEC in 1D systems and hence we do not expect BEC on poorly connected graphs. An interesting and open question is hence to characterise those graphs that are sufficiently well connected to exhibit this kind of phase transition. It is conjectured that graphs with “dimension 3 or more” exhibit this behaviour.

The aim of [1], however, is to show that this is not merely a phenomenon of the *free* Bose gas. Indeed the phase transition survives the introduction of hard-core repulsion of the particles. The main theorem gives even more detailed information as follows:

{thm:main}

Theorem 2.1. *The thermodynamic limit (1.4) exists for all values of $\rho \in [0, 1]$, $\beta \in [0, \infty)$ and is given by*

$$\beta f(\rho, \beta) = \rho \log \rho + (1 - \rho) \log(1 - \rho) + \beta \rho + \begin{cases} 0 & \text{if } 0 \leq \beta < \beta_c \\ \frac{(1-2\rho)^2}{2}(\beta - \beta_c) + \frac{1}{2} \log[1 - E(\beta)] + \rho(1 - \rho)\beta E(\beta) & \text{if } \beta_c \leq \beta \end{cases},$$

where $\beta_c = \beta_c(\rho) = \frac{1-\rho}{1-2\rho} \log \frac{1-\rho}{\rho} \in (0, \infty)$ and $E(\beta) \in [0, 1]$ is the unique solution of the equation

$$[(1 - 2\rho)^2 + 4\rho(1 - \rho)E]^{-\frac{1}{2}} \log \frac{1 + [(1 - 2\rho)^2 + 4\rho(1 - \rho)E]^{\frac{1}{2}}}{1 - [(1 - 2\rho)^2 + 4\rho(1 - \rho)E]^{\frac{1}{2}}} = \beta. \quad (2.3)$$

3. PROOF OF THE MAIN THEOREM

The outline of the proof is as follows:

- (1) Derive a random walk representation of Z using a Feynman-Kac type argument. (Lemma 3.1)
- (2) Notice that the relevant properties of the random walk representation can be kept track of in a 1D time-homogeneous Markov chain. (Lemma 3.3)
- (3) Use Ventsel's theorem to derive a large-deviation principle for the paths of the Markov chain — recall that we want “ $\frac{1}{V} \log Z_V$ ”. (Lemma 3.5)
- (4) Use the contraction principle to express the free energy in terms of a variational problem involving the rate function from before. (Lemma 3.9)
- (5) Analysing the variational problem using Euler-Lagrange calculus.

3.1. A random walk representation.

{lem:RWrepr}

Lemma 3.1. *Let $\mathbb{P}_{x_1, x_2, \dots, x_N}$ be the joint probability distribution of N independent continuous-time simple random walks $\eta_1(t), \dots, \eta_N(t)$ on the complete graph Λ_V , starting from the sites $x_1, \dots, x_N$, respectively. Denote by τ the first collision time of these random walks $\tau := \inf\{s \geq 0 \mid \eta_i(s) = \eta_j(s) \text{ for some } i, j \in \{1, \dots, N\}, i \neq j\}$. Then we have*

$$Z(\beta, N, V) = \binom{V}{N} \mathbb{P}_{1, \dots, N}(\tau > \beta \text{ and } \{\eta_1(\beta), \dots, \eta_N(\beta)\} = \{1, \dots, N\}). \quad (3.1) \quad \{\text{eq:RWrepr}\}$$

Proof. Since it is fairly standard the Feynman-Kac argument will only be sketched briefly: Write our Hamiltonian as $H_{N,V} = -\sum_{i=1}^N \Delta_i + \sum_{i < j}^N U(x_i - x_j)$. The trace of a matrix can be written as sum over its diagonal entries, so we need to find $A(\cdot, \cdot)$ such that $[e^{-\beta H_{N,V}} f](x) = \sum_{y \in \Lambda_V^N} A(x, y) f(y)$. To this end apply Trotter's product formula to get

$$[e^{-\beta H_{N,V}} f](x) = \lim_{n \rightarrow \infty} \left[\left(e^{\frac{\beta}{n} \sum_i \Delta_i} e^{-\frac{\beta}{n} \sum_{i < j} U(x_i - x_j)} \right)^n f \right](x).$$

Now note that $e^{-\frac{\beta}{n} \sum_{i < j} U(x_i - x_j)}$ still acts as a multiplication operator, while $e^{\frac{\beta}{n} \sum_i \Delta_i}$ is the generator of the simple random walk on Λ_V on a time interval of length β/n (independently on every coordinate x_i). Using this fact repeatedly we arrive at

$$[e^{-\beta H_{N,V}} f](x) = \sum_{y \in \Lambda_V^N} \underbrace{\left(\prod_{i=1}^N \int_{\eta_i: x_i \mapsto y_i} \right)}_{=A(x,y)} e^{-\sum_{i < j} \int_0^\beta U(\eta_i(s) - \eta_j(s)) ds} d\mathbb{P}_{x_1, \dots, x_N}(\eta_1, \dots, \eta_N) f(y).$$

It is easy to see that to calculate the trace over the *symmetric subspace* we have to replace the sum over the diagonal of A (this would correspond to considering N dependent SRWs conditioned to return to their origin at time β) by $\frac{1}{N!} \sum_{\sigma \in S_N} \sum_{x \in \Lambda_V^N} A(x, \sigma(x))$, where S_N is the set of permutations on N elements. This corresponds to considering N dependent SRWs conditioned to end at $\sigma(x)$ (where x is its starting point and $\sigma(x) := (x_{\sigma(1)}, \dots, x_{\sigma(N)})$).

Plugging in the definition of U we obtain

$$Z(\beta, N, V) = \frac{1}{N!} \sum_{x_1, \dots, x_N \in \Lambda_V} \sum_{\sigma \in S_N} \mathbb{P}_{x_1, \dots, x_N}(\tau > \beta \text{ and } \eta_i(\beta) = x_{\sigma(i)} \forall i),$$

which readily gives (3.1). $\square$

Remark 3.2. *At this point it might be tempting to proceed analysing (3.1) using a coloring argument like in the proof of Theorem 2 of [3]. Due to the fixed number of particles N , however, we lose independence of colorings of different sites. Up to different constants, due to the law of large numbers, one might still expect a similar result as in [3], but the analysis gets significantly more involved due to the lack of independence. The situation seems to be very similar to the analysis of the two definitions of Erdős-Rényi graphs: One with a fixed number of edges distributed in a uniformly random way (Erdős' original definition) or as a percolation problem on the complete graph; see e.g. [Wikipedia](#).*

Interestingly, this is not the only way cycles enter the analysis of this problem. See e.g. approaches using a different graphical representation (and representation theory) to extend the result of [1]: Theorem 1.2 of [4] and Theorem 1 of [5]. Here loops arise, as explained in [7], by applying the Lie-Trotter product formula to (4.3) and then taking the trace.

3.2. Reducing the problem to a Markov chain. Consider the continuous-time Markov chain $Y^{V,N}$ on $\mathbb{Z}_{\geq 0}$ with unit jump rate and transition rates as follows:

$$\begin{aligned} P_{k \rightarrow k+1}^{V,N} &= \frac{N-k}{N} \frac{V-N-k}{V-N} \\ &= 1 - \frac{V}{V-N} \frac{k}{N} + \frac{N}{V-N} \left(\frac{k}{N}\right)^2 =: r^{V,N} \left(\frac{k}{N}\right) \\ P_{k \rightarrow k-1}^{V,N} &= \frac{k}{N} \frac{k}{V-N} =: l^{V,N} \left(\frac{k}{N}\right) \\ P_{k \rightarrow k}^{V,N} &= \frac{k}{N} \frac{V-2k}{V-N} \\ &= 1 - r^{V,N} \left(\frac{k}{N}\right) - l^{V,N} \left(\frac{k}{N}\right) =: q^{V,N} \left(\frac{k}{N}\right). \end{aligned}$$

Lemma 3.3.

$$Z(\beta, N, V) = \binom{V}{N} \exp\left(-\frac{N(N-1)}{V}\beta\right) \mathbb{P}\left(Y^{V,N} \left(\frac{N(V-N)}{V}\beta\right) = 0\right). \quad (3.2)$$

Due to lack of space we do not present the proof here, but chose instead to discuss an extension of these techniques to different settings:

Remark 3.4 (Generalisation to non-trivial geometries). *Until Lemma 3.3 we did not make substantial use of the fact that we are considering the complete graph. In particular it is not hard to see that we should not expect to get a one-dimensional Markov chain encoding the dynamics needed to express (3.1). Clearly the killing rate of $\tilde{Y}^{V,N}$ will depend on $\tilde{Y}^{V,N}$ as soon as the number of particles that are nearest neighbours is not constant anymore. This problem has been addressed in [3] by choosing a pair potential that cancels*

these terms. The bigger problem, however, is to find a higher, but still finite-dimensional analogue of $\tilde{Y}^{V,N}$ that preserves the Markov property of $\mathbf{1}\{\{\eta_i(0)\}_i = \{\eta_i(s)\}_i\}$, $s \in [0, \beta]$.

The two requirements, being finite-dimensional and preserving the Markov property, are hard to fulfill simultaneously: On the one hand we can easily preserve the Markov property by taking a parameter set that grows in N by considering the $N \times N$ matrix K with $K_{ij} := d(x_i, x_j)$. This clearly characterises the particle configuration $\{\eta_i(s)\}_i$ up to isometry; in particular the transition probabilities exist (giving the desired Markov property) and can be given explicitly. As the dimension of the parameter set grows in V , however, it does not make sense to think of a scaling limit. On the other hand we could take $\tilde{Y}^{V,N}$ just as in [1]. But in general (i.e. on all graphs that are not the complete graph) this is not a Markov process anymore.

Intuitively we do not keep track of enough information; to see that it makes a difference of how we got to, say, $\tilde{Y}^{V,N} = 1$ consider the following two paths leading to this point: (1) We fluctuate between $\tilde{Y}^{V,N} = 0$ and $\tilde{Y}^{V,N} = 1$ and (2) since the first time step we have $\tilde{Y}^{V,N} = 1$. In case (1) the one particle that is not in $\{1, \dots, N\}$ has to have distance one from $\{1, \dots, N\}$ whereas in case (2) the distance of the particle not in $\{1, \dots, N\}$ is likely to be much larger after a sufficiently long time, hence there is a larger chance of not going to $\tilde{Y}^{V,N} = 0$ in the next step.

This suggests that families of growing graphs which are vertex-transitive and of uniformly finite diameter should be the most promising ones to extend this proof to. The easiest case might be the complete bipartite graph $K_{V,V}$ (and considerations should readily generalise to complete k -partite graphs). The “next easiest” case, the 2D Hamming graph, which also has diameter equal to 2, already seems intractable.* By the lack of (complete) symmetry of $K_{V,V}$ we have to consider all the different (up to isometry) ways the set $\{1, \dots, N\}$ can be arranged in $K_{V,V}$. To this end fix any $\delta \in [0, 1/2]$ and consider $\{1, \dots, [\delta N]\}$ in one cluster (call it left) of $K_{V,V}$ and $\{[\delta N] + 1, \dots, N\}$ in the other one (call it right). This covers all possible cases. Instead of $\tilde{Y}^{V,N}$ we consider for each combination of N, V, δ the Markov chain (C^L, C^R, R) where

$$C_t^L := |\eta_t \cap \{1, \dots, [\delta N]\}|, \quad C_t^R := |\eta_t \cap \{[\delta N] + 1, \dots, N\}|$$

are the numbers of “correctly placed” particles on the left and the right cluster, respectively. Introduce furthermore R_t , the number of particles in the right cluster. The analogue of $\tilde{Y}^{N,V} = 0$ here is $C^L = [\delta N], C^R = N - [\delta N]$. For example the probability of going from (C^L, C^R, R) to $(C^L + 1, C^R - 1, R - 1)$ is given by the probability of picking a particle from the right cluster times the (conditional) probability of picking one of those occupying some site $i \in \{[\delta N] + 1, \dots, N\}$ (given that we pick one from the right) times the probability of dropping it on some site $j \in \{1, \dots, [\delta N]\}$ that is not occupied by another particle. This

*Where “intractable” means that I was not able to find a uniformly finite-dimensional representation of the number of nearest neighbour particles; this seems to be the minimal information one needs to keep track of in order to make $\mathbf{1}\{\{\eta(0)\} = \{\eta(\beta)\}\}$ a Markov process for the reason outlined above.

gives

$$\mathbb{P}((C^L, C^R, R) \mapsto (C^L + 1, C^R - 1, R - 1)) = \frac{R}{N} \frac{C^R}{R} \frac{\delta N - C^R}{V - (N - R)}. \quad (3.3) \quad \{\text{eq:transProbK}\}$$

Probabilities for all the other outcomes $(C^L, C^R - 1, R - 1)$, $(C^L, C^R, R - 1)$, $(C^L + 1, C^R, R - 1)$ and the same with left and right interchanged follow easily by negating (i.e. replacing it with $1 - \cdot$) all combinations of the factors in (3.3) (and adjusting the last term accordingly).

Though tedious, one should be able to adapt the rest of the proof for this setting (and more generally for complete k -partite graphs) as well.

3.3. Ventsel's LDP for paths. For notational simplicity introduce $\varepsilon = N^{-1}$. Recall briefly the definition of LDPs: A sequence of measures μ_n on the measure space $(\mathfrak{X}, \mathfrak{F})$ satisfy a LDP with *rate function* I (i.e. a lower semi-continuous map $\mathfrak{X} \rightarrow [0, \infty]$) and speed n if for all $F \in \mathfrak{F}$:

$$\{\text{eq:LDPdef}\} \quad - \inf_{x \in F^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x \in F} I(x). \quad (3.4)$$

To obtain the free energy from (3.2) we still have to evaluate

$$\{\text{eq:LDPProb}\} \quad \frac{1}{V} \log \mathbb{P}(Y^{V,N}(N\theta) = 0) = \rho \frac{1}{N} \log \mathbb{P}(N^{-1}Y^{V,N}(N\theta) = 0), \quad (3.5)$$

with $\theta = (1 - \rho)\beta$. The right hand side suggests to treat it as a large deviation problem (note that we smuggled in a factor of N^{-1} to make the random paths ξ^ε given by

$$\xi^\varepsilon(t) := N^{-1}Y^{V,N}(Nt)$$

comparable, i.e. have the same order of magnitude and hence have a non-trivial limit). The sequence of measures to be considered will be the laws of ξ^ε on the set of cadlag* functions $D_0[0, \theta]$ equipped with its natural topology (i.e. the Skorohod topology; note that this reduces to the sup-norm topology on the subspace of continuous functions $C_0[0, \theta]$). Denote furthermore by $AC_0[0, \theta]$ the subspace of absolutely continuous functions on $[0, \theta]$. In this setting Ventsel [9] proved the following:

Lemma 3.5. *The processes $\xi^\varepsilon \in D_0[0, \theta]$ satisfy a LDP with rate function or “action functional”*

$$\{\text{ionFunctional}\} \quad \mathcal{I}_\rho[y] := \begin{cases} \int_0^\theta L(y(s), \dot{y}(s)) ds + \theta & \text{if } y \in AC_0[0, \theta] \text{ and the integral is finite} \\ \infty & \text{otherwise,} \end{cases} \quad (3.6)$$

where the Lagrangian L is given by

$$L(x, v) := \frac{v}{2} \log \left(\frac{\sqrt{v^2 + 4l(x)r(x)} + v \frac{l(x)}{r(x)}}{\sqrt{v^2 + 4l(x)r(x)} - v \frac{l(x)}{r(x)}} \right) - \sqrt{v^2 + 4l(x)r(x)} \underbrace{-1 + l(x) + r(x)}_{=-q(x)}$$

with the appropriately rescaled scaling limits of $l^{V,N}, r^{V,N}, q^{V,N}$ in the limit $\lim_{N, V \rightarrow \infty} \frac{N}{N/V \rightarrow \rho}$:

$$r(x) = 1 - \frac{1}{1 - \rho}x + \frac{\rho}{1 - \rho}x^2, \quad l(x) = \frac{\rho}{1 - \rho}x^2, \quad q(x) = 1 - l(x) - r(x).$$

*This is the natural setting for Poisson jump processes that we are considering.

Remark 3.6. *Note that the expression for the Lagrangian (plus the constant θ) does indeed coincide with the Laplacian given by Ventsel in [8].*

Since the proof is essentially (up to minor modifications) given in [9] we will not delve into it any further.

Remark 3.7 (Why do we look at paths?). *Thinking of the paths as collection of partial sums of the increments, one might naively wonder whether a LDP for the entire path (as opposed to just the value of the path at $t = \theta$) is actually needed. If l, r, q were constant, the increments would be identically distributed and classical LDP theory, e.g. Cramer's theorem, would suffice to get a LDP for the final value (i.e. "last partial sum").*

In this case (a spatially inhomogeneous Markov chain), however, the n -th increment depends on the $(n - 1)$ -th partial sum. In terms of increments, this introduces a lot of complicated correlations which should convince the sceptical reader of the need for using heavy machinery like Ventsel's result.

Remark 3.8 (A plausibility argument for Ventsel's result). *We want to push the analogy to Cramer's theorem a little bit more. To this end we (vaguely) state it in terms of increments: Let X_i be a collection of iid variables and $S_m := \sum_{i=1}^m X_i$ be their partial sums. Consider the log moment generating function $G(t) := \log \mathbb{E}(e^{tX_1})$ of the increments, say X_1 . Then the laws of $N^{-1}S_N$ ($N \rightarrow \infty$) satisfy a LDP with rate function $G^* = H$, the Legendre transform of G .*

Clearly we have $G(t) = \log \mathbb{E}(e^{t(S_{m+1}-S_m)})$ for all m . Now let the distributions of the increments X_{i+1} depend on i and S_i . Clearly Cramer's theorem does not apply anymore, but it suggests investigating the quantity $G(k, x; t) := \log \mathbb{E}_{k,x} e^{t(S_{k+1}-x)}$ and its Legendre transform (in the last variable only) $H(k, x; \cdot) = G^*(k, x; \cdot)$. Now write $N^{-1}S_{Nt} =: \phi(t)$ for N large and note that

$$\frac{1}{N} \sum_{i=Nt}^{N(t+\delta t)} X_i \stackrel{(def)}{=} \phi(t+\delta t) - \phi(t) = \phi'(t)\delta t + o(\delta t).$$

For suitably continuous paths (and suitable probability distributions of X_i changing not too much as i, S_i vary) Cramer's theorem hence gives that

$$\mathbb{P}_{k=Nt, \phi(t)=x} \left(\phi'(t) + o(1) = \frac{1}{N\delta t} \sum_{i=Nt}^{N(t+\delta t)} X_i \right) \approx e^{-N\delta t H(t, \phi(t); \phi'(t))}.$$

While the variables X_i may depend on S_{i-1} and i , they do not depend on the realisation of the path leading up to (i, S_i) by the Markov property. In particular conditioning on $N^{-1}S_i = x$ essentially makes them (conditionally) independent. But this is just what we need as we are considering $H(t, x; \cdot)$. Thus we conclude

$$\mathbb{P}(\phi = \varphi + o(1)) \approx \prod_{\delta t} e^{-N\delta t H(t, \varphi(t); \varphi'(t))} \approx e^{-N \int_0^1 H(t, \varphi(t); \varphi'(t)) dt}$$

to obtain a very crude heuristic derivation for Venttsel's theorem.*

3.4. An application of the contraction principle. Note that the right hand side of (3.5) is, up to the constant ρ just $\varepsilon \log \mathbb{P}(\xi^\varepsilon(\theta) = 0)$. By the contraction principle (the laws of) $\xi^\varepsilon(\theta)$ satisfy a LDP with rate function

$$\{eq: defRate\} \quad I_\rho(\theta) := \inf \{ \mathcal{I}_\rho[y] \mid y \in C_0[0, \theta] : y(0) = 0 = y(\theta) \}. \quad (3.7)$$

Recalling the definition of a LDP (see (3.4)), this is barely not enough to prove the following lemma:

Lemma 3.9. *The limit $\lim_\varepsilon \varepsilon \log \mathbb{P}(\xi^\varepsilon(\theta) = 0)$ exists and is equal to $I_\rho(\theta)$.*

Sketch of the proof. It is enough to show that $-I_\rho(\theta) \leq \varepsilon \log \mathbb{P}(\xi^\varepsilon(\theta) = 0) \leq -I_\rho(\theta)$. The upper bound follows from the above remark and (3.4) since $F = \{y : y(0) = 0 = y(\theta)\}$ is already closed.

The lower bound is more complicated and technical. The main issue is that, since we want the reverse direction (and all we can use at this point is the LDP), we will have to work with some open set instead of F . It remains to show that we can approximate F by $F_\delta := \{y : y(0) = 0, y(\theta - \delta) < \delta\}$ at a “sufficiently low price”. Due to lack of space we will not present the full proof here, but note that this is a delicate issue involving some phase space analysis of the solutions to the corresponding Euler-Lagrange equation. $\square$

3.5. Evaluating the rate function. We want to evaluate (3.7). To this end, note that every local minimum y in $C_0[0, \theta]$ satisfies the Euler-Lagrange equation:

$$\{eq: EulerL\} \quad \frac{\partial L}{\partial x} L(y, \dot{y}) = \frac{d}{dt} \frac{\partial L}{\partial v}(y, \dot{y}). \quad (3.8)$$

We may thus restrict our attention to solutions of (3.8). Since the Lagrangian L is fairly complicated an explicit solution seems intractable. Note, however, that the Lagrangian L is invariant under time shifts. Noether's theorem thus tells us that $v \frac{\partial L}{\partial v}(x, v) - L(x, v) = \sqrt{v^2 + 4l(x)r(x)} + q(x)$ is constant *along trajectories* y satisfying (3.8). For every such y denote its “energy” $\sqrt{\dot{y}(t)^2 + 4l(y(t))r(y(t))} + q(y(t)) \equiv E \in [0, \infty)$. We conclude that to at least one $E \in [0, \infty)$ there corresponds at least one solution $y_E \in C_0[0, \theta]$ of the Euler-Lagrange equation.

Solving $\sqrt{v^2 + 4l(x)r(x)} + q(x) = E$ for v^2 (using the definitions of l, r, q) gives $v^2 = a_E(x - b_E)^2 - c_E$ for constants

$$\begin{aligned} a_E &= \frac{(1 - 2\rho)^2 + 4\rho(1 - \rho)E}{(1 - \rho)^2} \\ b_E &= \frac{(1 - \rho)E}{(1 - 2\rho)^2 + 4\rho(1 - \rho)E} \\ c_E &= \frac{4\rho(1 - \rho)E^2(1 - E)}{(1 - 2\rho)^2 + 4\rho(1 - \rho)E}. \end{aligned}$$

* H is used in Venttsel's paper, here it actually corresponds to our Lagrangian L . For a technically less (but still) daunting result that includes the key ideas outlined above, see Mogulskii's theorem.

This way we obtain for every $E \in [0, \infty)$ two differential equations

$$\text{eq:EulerL_ODE}\} \quad \dot{y}_E^\pm(t) = \pm \sqrt{a_E(y_E^\pm(t) - b_E)^2 - c_E} \quad (3.9)$$

giving two solutions to the Euler-Lagrange equation at most one of which will return to zero. Denote this one by y_E (choose arbitrarily if neither return to 0). It can be checked that for $E \in [1, \infty)$ the solutions to these ODEs (with initial condition $y_E(0) = 0$) do not return to 0. In particular they will not satisfy $y_E(\theta) = 0$ so we may restrict our attention to solutions of (3.9) with $E \in (0, 1)$.

It can be seen that they go up to $\sup_t y_E(t) =: \bar{x}_E = b_E - \sqrt{\frac{c_E}{a_E}}$ and then return to 0 at some time

$$\hat{\theta}(E) := \inf\{s > 0 : y_E(s) = 0\}.$$

Denote by $\hat{S}(E)$ the action of this trajectory, i.e.

$$\hat{S}(E) := \int_0^{\hat{\theta}(E)} L(y_E(s), \dot{y}_E(s)) ds.$$

First we want to find all $E = E(\theta) \in (0, 1)$ such that $\hat{\theta}(E) = \theta$. Writing $\hat{\theta}(E) = 2 \int_0^{\bar{x}_E} \frac{dx}{\dot{y}_E(x)}$ we obtain

$$\hat{\theta}(E) = \frac{1 - \rho}{\sqrt{(1 - 2\rho)^2 + 4\rho(1 - \rho)E}} \log \frac{1 + \sqrt{(1 - 2\rho)^2 + 4\rho(1 - \rho)E}}{1 - \sqrt{(1 - 2\rho)^2 + 4\rho(1 - \rho)E}}.$$

Note that $(0, 1) \ni E \mapsto \hat{\theta}(E) \in \mathbb{R}_+$ is a C_∞ monotonically increasing function with

$$\hat{\theta}(0) = \frac{1 - \rho}{1 - 2\rho} \log \frac{1 - \rho}{\rho} =: \theta_c$$

and

$$\lim_{E \rightarrow 1} \hat{\theta}(E) = \infty.$$

Hence, for any $\theta \in [\theta_c, \infty)$ there is a unique $E(\theta) \in [0, 1)$ satisfying

$$\hat{\theta}(E) = \theta, \quad (3.10) \quad \{\text{eq:thetaE}\}$$

while for $\theta \in [0, \theta_c)$ (3.10) has no solution. This just means that for $\theta \in (\theta_c, \infty)$ the Euler-Lagrange equation with initial and final conditions $y(0) = y(\theta) = 0$ has two solutions: (1) $y_0(s) \equiv 0$, (2) $y_{E(\theta)}$, where the second one minimises the action functional in (3.7). For $\theta \in [0, \theta_c)$ the only solution is (1). Noting that $L(0, 0) = 0$ we thus arrive at

$$I_\rho(\theta) - \theta = \begin{cases} 0 & \text{if } \theta \in [0, \theta_c] \\ \hat{S}(E(\theta)) & \text{if } \theta \in (\theta_c, \infty). \end{cases}$$

It remains to simplify $S(E(\theta))$. This calculation is carried out in the paper. The essential steps are as follows:

- Realise that using the conservation law (3.9) and the definition of L we can write $\hat{S}(E)$ as

$$\hat{S}(E) = \int_0^{\bar{x}_E} \log \frac{E - q(x) + v_E(x)}{E - q(x) - v_E(x)} dx - E\hat{\theta}(E).$$

- Calculate the derivative of $\hat{S}(E)$ with respect to E to obtain

$$\frac{d\hat{S}}{dE}(E) = -E \frac{d\hat{\theta}}{dE}(E).$$

- Combining the above two steps we get

$$\hat{S}(E(\theta)) = - \int_{\theta_c}^{\theta} E(\theta') d\theta'$$

for $\theta \in (\theta_c, \infty)$.

- After integration by parts one finally concludes the desired result:

$$I_\rho(\theta) - \theta = \begin{cases} 0 & \text{if } \theta \in [0, \theta_c] \\ \frac{(1-2\rho)^2}{2(1-\rho)\rho}(\theta - \theta_c) + \frac{1}{2\rho} \log(1 - E(\theta)) + \theta E(\theta) & \text{if } \theta \in (\theta_c, \infty). \end{cases}$$

4. IMPLICATIONS OF THE RESULT

While an exact formula for the free energy of some system certainly has many implications for this system, we want to focus on implications for the spin- $\frac{1}{2}$ quantum Heisenberg ferromagnet (QHF). To the novice the connection is not obvious, but tacitly assumed to be known in the literature (see e.g. [4] (Theorem 1.2), [5] (Theorem 1) for an incomplete list of extensions of [1]) so let us sketch it briefly:

4.1. Equivalence of SSEP and QHF. Fix some set Λ , $S = 1/2$ (note that the construction below works for general $S \in \frac{1}{2}\mathbb{N}$) and consider the Hilbert space

$$\mathcal{H}_\Lambda := \bigotimes_{x \in \Lambda} \mathbb{C}^2 \simeq \text{span}\{(\pm 1)^\Lambda\},$$

with Hamiltonian

$$H_\Lambda = J \sum_{x \sim y} S_x^1 S_y^1 + S_x^2 S_y^2 + S_x^3 S_y^3.$$

Here $\{S_x^i\}_{x \in \Lambda}^{i=1,2,3}$ are hermitian matrices on $\mathbb{C}^{2S+1} = \mathbb{C}^{2*}$ satisfying the following properties for all $x \in \Lambda$:

$$\begin{aligned} \{\text{eq:Sreqs}\} \quad & [S_x^1, S_x^2] = iS_x^3, \quad [S_x^2, S_x^3] = iS_x^1, \quad [S_x^3, S_x^1] = iS_x^2, \\ & (S_x^1)^2 + (S_x^2)^2 + (S_x^3)^2 = S(S+1)id = \frac{3}{4}id. \end{aligned} \tag{4.1}$$

Existence follows by construction: Let $|a\rangle$, $a \in \{\pm \frac{1}{2}\}$ denote an orthonormal basis of $\mathbb{C}^{2S+1} = \mathbb{C}^2$. Set $S_x^3|a\rangle := a|a\rangle$ and define $S_x^\pm$ by

$$S_x^\pm |a\rangle := \sqrt{S(S+1) - a(a \pm 1)} |a \pm 1\rangle.$$

After setting $S_x^1 = \frac{1}{2}(S_x^+ + S_x^-)$ and $S_x^2 = \frac{1}{2i}(S_x^+ - S_x^-)$ it can easily be checked that the relations in (4.1) are satisfied.

*It is easier to introduce them on $\mathbb{C}^2$, but by abuse of notation they act on $\mathcal{H}_\Lambda$ as $S_x^i \otimes id_{\Lambda \setminus \{x\}}$, where $id_{\Lambda \setminus \{x\}}$ is the identity in $\bigotimes_{y \in \Lambda \setminus \{x\}} \mathbb{C}^2$.

Note that $S_x^\pm$ correspond to bosonic creation and annihilation operators a_x, a_x^* , respectively. An quick calculation furthermore reveals that

$$n_x := a_x^* a_x = S_x^+ S_x^- = S_x^3 + \frac{1}{2},$$

so the expected total number of particles $\sum_{x \in \Lambda} \langle n_x \rangle$ divided by $|\Lambda|$ (this is equal to ρ in the setting of [1]) is equal to $\frac{1}{2} + m$, where $m := \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \langle S_x^3 \rangle$ is the average magnetisation.

Note that the measure with respect to which we are taking expectations has not been defined yet. For observables A expressed in terms of spin operators the natural measure is $\langle A \rangle := \frac{\text{Tr} A e^{-\beta H_\Lambda}}{\text{Tr} e^{-\beta H_\Lambda}}$. In order to get that $\rho = \frac{1}{2} + m$ we need to show that the generator of the Symmetric Simple Exclusion Process (SSEP) L_Λ is unitarily equivalent to H_Λ (in that case the expectations will be the same). The following exposition will closely follow [6].

From (1.1) it is not hard to see that $H_{V,N}$ (our L_Λ) actually admits the following (more concise, but less explicit) representation:

$$Lf(\eta) = \sum_{x \sim y} c(x, y) (f(\eta) - f(\eta^{xy})),$$

where $\eta : \Lambda_V \rightarrow \{0, 1\}$ with $\sum_{x \in \Lambda_V} \eta_x = N$ indicates whether a particle is present at site $x \in \Lambda_V$ and η^{xy} is the configuration η with the values at x and y interchanged. We consider now the explicit isomorphism

$$f \mapsto \psi := \sum_{\eta} f(\eta) |\eta\rangle.$$

Clearly we have $S_x^3 |\eta\rangle = (\eta_x - \frac{1}{2}) |\eta\rangle$ and it is also not hard to see that

$$\frac{1}{2} - 2(S_x^1 S_y^1 + S_x^2 S_y^2 + S_x^3 S_y^3) = 1 - T_{xy}, \quad (4.2) \quad \{\text{eq:Txy}\}$$

where $T_{xy}|a, b\rangle := |b, a\rangle$. To see that under the isomorphism given above SSEP becomes the QHF we calculate its action:

$$\begin{aligned} Lf(\eta) &\mapsto \sum_{\eta} Lf(\eta) |\eta\rangle \\ &= \sum_{x \sim y} \sum_{\eta} c(x, y) (f(\eta) - f(\eta^{xy})) |\eta\rangle \\ &= \sum_{x \sim y} \sum_{\eta} c(x, y) (1 - T_{xy}) f(\eta) |\eta\rangle \\ &\stackrel{(4.2)}{=} H\psi, \end{aligned} \quad (4.3) \quad \{\text{eq:TxySSEP}\}$$

with $J = 2c$ (if $c(x, y) \equiv \text{const}$), up to an additive constant in the Hamiltonian (which is sometimes added).

4.2. SSEP $\leftrightarrow$ QHF correspondences.

- To the (conserved) number of particles $N = \rho V$ in SSEP corresponds the (conserved) sum of spins in direction 3, i.e. the total magnetisation.
- To the unique invariant measure of SSEP (the uniform measure) corresponds the ferromagnetic ground state with magnetisation $N - V/2$.
- The spin flip symmetry ($|\frac{1}{2}\rangle \rightarrow -|\frac{1}{2}\rangle$) corresponds to $N \mapsto V - N$, a slightly more subtle symmetry in the SSEP picture which can also be seen directly as follows: Since we are dealing with indistinguishable particles we can only “see swaps of particle with no particle” (see (4.3) and the definition of T_{xy}). In particular the roles of particle and no particle are completely symmetric; this symmetry would be lost if we were allowed to distinguish particles as one could then swap particles, but swapping “no particles” doesn’t make sense.

REFERENCES

- [1] Tóth, B. *Phase transition in an interacting Bose system. An application of the theory of Ventsel and Freidlin*. Journal of statistical physics 61.3-4 (1990): 749-764.
- [2] Penrose, O. *Bose-Einstein condensation in an exactly soluble system of interacting particles*. Journal of Statistical Physics 63.3-4 (1991): 761-781.
- [3] Tóth, B. *Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet*. Letters in Mathematical Physics 28.1 (1993): 75-84.
- [4] Björnberg, J. *The free energy in a class of quantum spin systems and interchange processes*. Journal of Mathematical Physics 57.7 (2016): 073303.
- [5] Alon, G. and Kozma, G. *The mean-field quantum Heisenberg ferromagnet via representation theory*. arXiv preprint arXiv:1811.10530 (2018).
- [6] Nachtergaele, B. *Quantum spin systems after DLS 1978*. Proceedings of Symposia in Pure Mathematics. Vol. 76. No. 1. Providence, RI; American Mathematical Society; 1998, 2007.
- [7] Aizenman, M. and Nachtergaele, B. *Geometric aspects of quantum spin states*. Communications in Mathematical Physics 164.1 (1994): 17-63.
- [8] Ventsel, A. *Rough limit theorems on large deviations for Markov stochastic processes. I*. Theory of Probability & Its Applications 21.2 (1977): 227-242.
- [9] Ventsel, A. *Rough limit theorems on large deviations for Markov stochastic processes. III*. Theory of Probability & Its Applications 24.4 (1980): 675-692.