Large Deviations for Ablowitz-Ladik lattice, and the Schur flow

Guido Mazzuca, Ronan Memin

Large Deviations for Ablowitz-Ladik lattice, and the Schur flow

G. Mazzuca*, R. Memin†

January 23, 2023

Abstract

We consider the Generalized Gibbs ensemble of the Ablowitz-Ladik lattice, and the Schur flow. We derive large deviations principles for the distribution of the empirical measures of the equilibrium measures for these ensembles. As a consequence, we deduce their almost sure convergence. Moreover, we are able to characterize their limit in terms of the equilibrium measure of the Circular, and the Jacobi beta ensemble respectively.

1 Introduction

The defocusing Ablowitz-Ladik (AL) lattice is the system of ODEs

$$i\dot{\alpha}_j = -(\alpha_{j+1} + \alpha_{j-1} - 2\alpha_j) + |\alpha_j|^2(\alpha_{j-1} + \alpha_{j+1})\,, \tag{1.1}$$

eq:AL

that describe the evolution of the complex functions $\alpha_j(t)$, $j \in \mathbb{Z}$ and $t \in \mathbb{R}$, here $\dot{\alpha}_j = \dfrac{d\alpha_j}{dt}$. We assume N-periodic boundary conditions $\alpha_{j+N} = \alpha_j$, for all $j \in \mathbb{Z}$. For simplicity, we consider the case N even, and, when not mentioned, the limits as $N \to \infty$ is taken along N even. This system was introduced by Ablowitz and Ladik [2,3] as a spatial discretization of the defocusing Nonlinear Schrödinger Equation (NLS)

$$i\partial_t\psi(x,t) = -\frac{1}{2}\partial_x^2\psi(x,t) + |\psi(x,t)|^2\psi(x,t).$$

The NLS is a well-known integrable model [35], and the Ablowitz-Ladik lattice is one of the several discretizations that preserve integrability [26].

It is straightforward to verify that the two quantities

$$K^{(0)} := \prod_{j=1}^{N}\left(1 - |\alpha_j|^2\right), \quad K^{(1)} := -\sum_{j=1}^{N}\alpha_j\overline{\alpha}_{j+1},$$

eq:K0_K1

are constants of motion for the AL lattice. Since $K^{(0)}$ is conserved along the flow, it implies that if $|\alpha_j(0)| < 1$ for all $j = 1, \dots, N$, then $|\alpha_j(t)| < 1$ for all times. Thus, we can consider $\mathbb{D}^N$ as our phase space, where $\mathbb{D} = \{z \in \mathbb{C}\,|\,|z| < 1\}$.

On this phase space we consider the symplectic form ω [14,17]

$$\omega = i\sum_{j=1}^{N}\frac{1}{\rho_j^2}\mathrm{d}\alpha_j \wedge \mathrm{d}\overline{\alpha}_j\,, \quad \rho_j = \sqrt{1 - |\alpha_j|^2}\,.$$

eq:symplect

*Department of Mathematics, The Royal Institute of Technology, Stockholm, Sweden.
email: mazzuca@kth.se
†UMPA ENS de Lyon, Lyon, France.
email: ronan.memin@ens-lyon.fr

The corresponding Poisson bracket is defined for functions $f, g \in \mathcal{C}^\infty(\mathbb{D}^N)$ as

$$\{f,g\} = i \sum_{j=1}^{N} \rho_j^2 \left(\frac{\partial f}{\partial \overline{\alpha}_j} \frac{\partial g}{\partial \alpha_j} - \frac{\partial f}{\partial \alpha_j} \frac{\partial g}{\partial \overline{\alpha}_j} \right). \tag{1.2}$$

eq:poisson_

Using this Poisson bracket, it is possible to rewrite the equations of motion (1.1) of the AL lattice in Hamiltonian form as

$$\dot{\alpha}_j = \{\alpha_j, H_{AL}\}, \quad H_{AL}(\boldsymbol{\alpha}) = -2\ln(K^{(0)}) + K^{(1)} + \overline{K^{(1)}}\,,$$

eq:hamilton

here $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$.

Conserved quantities. As we already mentioned, the AL lattice is an integrable model: this was proved by Ablowitz and Ladik [1,2]. Specifically, they were able to obtain a Lax pair for the Ablowitz-Ladik lattice by discretizing the 2×2 Zakharov-Shabat Lax pair of the cubic nonlinear Schrödinger equation.

Nenciu and Simon in [25, 27] constructed a new Lax pair for the Ablowitz-Ladik lattice, exploiting the connection of this system to the orthogonal polynomials on the unit circle. This link is the analogue of the well-known link between the Toda lattice and orthogonal polynomials on the real line (see e.g. [8]). This connection was also generalized to the non-commutative case [6].

Following [25,27], we construct the Lax matrix as follows. Consider the 2×2 unitary matrices

$$\Xi(\alpha_j) = \Xi_j = \begin{pmatrix} \overline{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}, \quad \rho_j = \sqrt{1-|\alpha_j|^2}, \quad j = 1, \ldots, N\,,$$

matrix_Xi

and the $N \times N$ matrices

$$\mathcal{M} = \begin{pmatrix} -\alpha_N & & & & & \rho_N \\ & \Xi_2 & & & & \\ & & \Xi_4 & & & \\ & & & \ddots & & \\ & & & & \Xi_{N-2} & \\ \rho_N & & & & & \overline{\alpha}_N \end{pmatrix}, \qquad \mathcal{L} = \begin{pmatrix} \Xi_1 & & & \\ & \Xi_3 & & \\ & & \ddots & \\ & & & \Xi_{N-1} \end{pmatrix}. \tag{1.3}$$

eq:cM_cL

Now let us define the Lax matrix

$$\mathcal{E} = \mathcal{L}\mathcal{M}\,, \tag{1.4}$$

eq:Lax_matr

which has the following structure

$$\mathcal{E} = \begin{pmatrix} * & * & * & & & & & & & * \\ * & * & * & & & & & & & * \\ & * & * & * & * & & & & & \\ & * & * & * & * & & & & & \\ & & & & & \ddots & \ddots & & & \\ & & & & & & * & * & * & * \\ & & & & & & * & * & * & * \\ * & & & & & & & * & * & * \\ * & & & & & & & * & * & * \end{pmatrix}.$$

The matrix $\mathcal{E}$ is a periodic CMV matrix (after Cantero, Moral and Velazquez [7]). It is straightforward to verify that the equations of motions (1.1) are equivalent to the following Lax equation for the matrix $\mathcal{E}$:

$$\dot{\mathcal{E}} = i\left[\mathcal{E}, \mathcal{E}^+ + (\mathcal{E}^+)^\dagger\right], \tag{1.5}$$

eq:Lax_pair

where † stands for hermitian conjugate, and

$$\mathcal{E}^+_{j,k} = \begin{cases} \frac{1}{2}\mathcal{E}_{j,j} & j = k \\ \mathcal{E}_{j,k} & k = j+1 \bmod N \text{ or } k = j+2 \bmod N \\ 0 & \text{otherwise.} \end{cases}$$

rem:unitary **Remark 1.1.** *We notice that since all the Ξ_j are unitary, then also $\mathcal{E}$ is unitary, this implies that all the eigenvalues λ_j lie on the unit circle, and they can be written in terms of their argument, namely for all $j = 1, \dots, N$ there exists a $\theta_j \in \mathbb{T} := [-\pi, \pi)$ such that*

$$\lambda_j = e^{i\theta_j} .$$

In view of this identification, and in order to simplify the notations, for any function $f(z) : \partial\mathbb{D} \to \mathbb{R}$, we write $f(\theta)$ in place of $f(e^{i\theta})$ when it is convenient. Further, we will write indifferently $\int_{\mathbb{T}} f(\theta)\mathrm{d}\mu(\theta)$ or $\int_{\partial\mathbb{D}} f(z)\mathrm{d}\mu(z)$ for any probability measure μ having support on the circle.

Remark 1.2. *We notice that $(\mathcal{E}^+)^\dagger + (\mathcal{E}^\dagger)^+ = \mathcal{E}^\dagger$ and $[\mathcal{E}, \mathcal{E}^\dagger] = 0$ since $\mathcal{E}$ is unitary. Therefore, the Lax pair (1.5) can be rewritten in the equivalent form*

$$\dot{\mathcal{E}} = i\left[\mathcal{E}, \mathcal{E}^+ - (\mathcal{E}^\dagger)^+\right] . \qquad \texttt{eq:Lax\_pair}$$

The formulation (1.5) implies that the quantities

$$K^{(\ell)} = \mathrm{Tr}\left(\mathcal{E}^\ell\right), \quad \ell = 1, \dots, N-1, \qquad \texttt{eq:constant}$$

are constants of motion for the defocusing AL system (1.1).

As in [24, 28], we introduce the Generalized Gibbs ensemble for the Ablowitz-Ladik lattice, namely the following probability measure on the phase space $\mathbb{D}^N$

$$\mathrm{d}\mathbb{P}^{V,\beta}_{AL,N}(\alpha_1, \dots, \alpha_N) = \frac{1}{Z^{AL}_N(V,\beta)} \prod_{j=1}^N (1-|\alpha_j|^2)^{\beta-1} \mathbf{1}_{\{\alpha_j \in \mathbb{D}\}} \exp(-\mathrm{Tr}(V(\mathcal{E})))\mathrm{d}^2\boldsymbol{\alpha}, \qquad (1.6) \quad \texttt{GGE AL}$$

where $V(e^{i\theta}) : \mathbb{T} \to \mathbb{R}$ is a continuous function, $\mathbf{1}_A$ is the indicator function of the set A, and $Z^{AL}_N(V,\beta)$ is the partition function of the system

$$Z^{AL}_N(V,\beta) = \int_{\mathbb{D}^N} \prod_{j=1}^N (1-|\alpha_j|^2)^{\beta-1} \exp(-\mathrm{Tr}(V(\mathcal{E})))\mathrm{d}^2\boldsymbol{\alpha}.$$

Furthermore, we consider the empirical measure $\mu_N(\mathcal{E})$ of the eigenvalues $e^{i\theta_1}, \dots, e^{i\theta_N}$ of the matrix $\mathcal{E}$ (1.4), namely

$$\mu_N(\mathcal{E}) = \frac{1}{N}\sum_{j=1}^N \delta_{e^{i\theta_j(\mathcal{E})}} ,$$

here δ_x is the delta function centred at x, furthermore, we notice that we can just consider the arguments $\theta_1, \dots, \theta_N$ of the eigenvalues since the matrix $\mathcal{E}$ is unitary, see Remark 1.1.

Our main result is a large deviations principle (LDP) with good rate function for the sequence $(\mu_N(\mathcal{E}))$ under the law $\mathbb{P}^{V,\beta}_{AL,N}$ (1.6). Namely, denoting by $\mathcal{P}(\mathbb{T})$ the set of probability measures on the Torus $\mathbb{T}$ endowed with the topology of weak convergence, there exists a function $J^V_\beta : \mathcal{P}(\mathbb{T}) \to [0, +\infty]$ such that:

1. it is *lower semicontinuous/good*, namely for any $a \geqslant 0$, $\{\mu \in \mathcal{P}(\mathbb{T}) \mid J_\beta^V(\mu) \leqslant a\} \subset \mathcal{P}(\mathbb{T})$ is compact,

2. it satisfies a *large deviations lower bound*, namely for all $O \subset \mathcal{P}(\mathbb{T})$ open,

$$-\inf_O J \leqslant \liminf_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}(\mu_N(\mathcal{E}) \in O), \qquad \texttt{ineq:large d}$$

3. it satisfies a *large deviations upper bound*, namely for all $F \subset \mathcal{P}(\mathbb{T})$ closed,

$$\limsup_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}(\mu_N(\mathcal{E}) \in F) \leqslant -\inf_F J. \qquad \texttt{ineq:large d}$$

We refer to [9] for a general introduction to large deviations.

s strong ldp **Remark 1.3.** *We notice that, by compactness of $\mathcal{P}(\mathbb{T})$, it is sufficient to prove a **weak** large deviations principle, see [9, Section 1.2], which is equivalent to a full large deviations principle, except that the large deviation upper bound (point 3) holds only for **compact** subsets of $\mathcal{P}(\mathbb{T})$.*

From this large deviations principle we are able to deduce that $\mu_N(\mathcal{E})$ converges almost surely as N goes to infinity.

THM:MAIN_1 **Theorem 1.4.** *Let $\beta > 0$. For any continuous function $V : \mathbb{T} \to \mathbb{R}$ the following holds.*

a. the sequence $\mu_N(\mathcal{E})$ under the law $\mathbb{P}_{AL,N}^{V,\beta}$ satisfies a large deviations principle at speed N with a good rate function J_β^V,

b. J_β^V achieves its minimum at a unique probability measure ν_β^V,

c. $\mu_N(\mathcal{E})$ converges almost surely and in $L^1(\mathbb{T})$ towards ν_β^V.

Moreover, following [19, 24, 28], we are able to characterize the measure ν_β^V in terms of the equilibrium measure of the Circular beta ensemble at high temperature [20, 21]. More precisely, consider the functional $\mu \mapsto f_\beta^V(\mu)$ given, for any $\mu \in \mathcal{P}(\mathbb{T})$ absolutely continuous with respect to Lebesgue measure and with density $\frac{\mathrm{d}\mu}{\mathrm{d}\theta}$, by

$$f_\beta^V(\mu) = -\beta \int_{\mathbb{T}\times\mathbb{T}} \log\left(|e^{i\theta} - e^{i\varphi}|\right) \mu(\mathrm{d}\theta)\mu(\mathrm{d}\varphi) + \beta \log(2) + \int_{\mathbb{T}} V(\theta)\mu(\mathrm{d}\theta) + \int_{\mathbb{T}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta)\right) \mu(\mathrm{d}\theta) + \log(2\pi)\,.$$

It is shown in [20], that the previous functional reaches its minimum for a unique absolutely continuous probability measure μ_β^V. Moreover in [24] it is proved that this measure is almost surely differentiable with respect to β. Exploiting this result, and Theorem 1.4 we are able to show that

AL_RELATION **Theorem 1.5.** *For any continuous $V, f : \mathbb{T} \to \mathbb{R}$*

$$\int_{\mathbb{T}} f(\theta)\mathrm{d}\nu_\beta^V(\theta) = \partial_\beta \left(\beta \int_{\mathbb{T}} f(\theta)\mathrm{d}\mu_\beta^V(\theta)\right).$$

Thus, we obtain a unique characterization of the measure ν_β^V.

In [24, 28], the authors considered the GGE (1.6) with polynomial potential, and they were able to prove Theorem (1.5) for this particular class of potentials using a transfer operator technique. In this sense, we generalize their result, extending it to the class of continuous potentials.

In the last part of the manuscript, we consider another integrable model related to the Ablowitz-Ladik lattice, namely the so-called Schur flow [18]. Also for this system, the Lax matrix is $\mathcal{E}$ (1.4). Following the same construction as in the Ablowitz-Ladik lattice case, we define a GGE for this model. We are able to show analogous results to Theorem 1.4 and Theorem 1.5 for the Schur flow. The main difference is that in place of the Circular beta ensemble, we have the Jacobi one.

To give a wider overview of the relevant literature, we mention that, H. Spohn in [28], applying the theory of generalized hydrodynamics [11], argues that the correlation functions of the Ablowitz-Ladik lattice with respect to the GGE (1.6) show a ballistic behaviour. As we already mentioned, in [24] the authors rigorously proved Theorem 1.5 for polynomial potential $V(z)$. Moreover, they computed explicitly the density of states in the case $V(z) = \eta(z + \overline{z})$, which corresponds to the classical Gibbs ensemble.

Lastly, it is worth to mention that this link between random matrix and integrable system was first noticed in [31]. In this paper the author considered the GGE for the Toda lattice, and he was able to study this ensemble, comparing it with the Gaussian beta ensemble [12]. We refer to [19, 23, 24, 29, 30, 32] for subsequent developments.

In particular, our work was inspired by the recent paper [19]. In this paper, the authors obtained a large deviations principle for the Toda lattice, and obtain an analogous result to Theorem 1.5, where in place of the Circular beta ensemble, they had the Gaussian one.

The structure of the paper is the following. In Section 2, we prove the first point of Theorem 1.4. In Section 3, we collect some known results related to the Circular beta ensemble in the high-temperature regime. Moreover, we reformulate the already known large deviations principle for this ensemble in terms of the AL lattice. In Section 4, we conclude the proof of Theorem 1.4, and we prove Theorem 1.5. Section 5, is dedicated to the Schur flow, where we prove the analogue of Theorem 1.4 and Theorem 1.5 for this integrable model. Finally, we defer to the appendix the most technical results of our manuscript.

2 Existence of a Large deviations principle for the empirical measure of the Ablowitz-Ladik lattice

sec:weak_LPD

The aim of this section is to prove the first point of Theorem 1.4, namely to show that, for $\mathbb{N} \ni N \geqslant 2$ and even, the sequence of empirical measures $\mu_N(\mathcal{E}) = \frac{1}{N}\sum_{j=1}^{N} \delta_{e^{i\theta_j(\mathcal{E})}}$ satisfies a large deviations principle. The strategy of proof is the following. First, we show that if $\mathcal{E}$ is distributed according to $\mathbb{P}^{\beta}_{AL,N} := \mathbb{P}^{0,\beta}_{AL,N}$ defined in (1.6), then the sequence of random probability measures $(\mu_N(\mathcal{E}))_{N \text{ even}}$ satisfies a large deviations principle in $\mathcal{P}(\mathbb{T})$, the space of probability measures on $\mathbb{T}$, endowed with the topology of weak convergence. Since according to this toppology $\mathcal{P}(\mathbb{T})$ is compact, it suffices to show that the sequence $(\mu_N(\mathcal{E}))_{N \text{ even}}$ satisfies a **weak** large deviations principle, see Remark 1.3. Then, applying Varadhan's Lemma [13, Theorem 1.2.1], we obtain the existence of a large deviations principle for arbitrary continuous V, *i.e.* the first point of Theorem 1.4.

We also notice that when $V = 0$ in (1.6) the α_i's are independent and identically distributed (*i.i.d*) with distribution $\Theta_{2\beta+1}$, where Θ_ν is defined for $\nu > 1$ as the random variable such that for $f : \mathbb{C} \to \mathbb{R}$ bounded and measurable

$$\mathbb{E}[f(X)] = \frac{\nu - 1}{2\pi} \int_{\mathbb{D}} f(z)(1 - |z|^2)^{\frac{\nu-3}{2}} \mathrm{d}^2 z. \tag{2.1}$$

def theta

erpretation

Remark 2.1. *We recall that for integer $\nu \geqslant 2$, such measure has the following geometrical interpretation: if $\mathbf{u} = (u_1, \dots, u_{\nu+1})$ is chosen at random according to the surface measure on the unit sphere S^ν in $\mathbb{R}^{\nu+1}$, then $u_1 + iu_2$ is Θ_ν distributed [22].*

To show that the sequence $(\mu_N(\mathcal{E}))_{N \text{ even}}$ satisfies a weak large deviations principle according to the law $\mathbb{P}^{\beta}_{AL,N}$, we only need the α_i's to be *i.i.d* according to some law σ with $\text{supp}(\sigma) \subseteq \overline{\mathbb{D}}$. Thus, we just assume the latter hypothesis, and we prove the result in more generality.

2.1 Large Deviations Principle for periodic CMV matrix

Let d be the distance on $\mathcal{P}(\mathbb{T})$ defined by

$$d(\mu,\nu) = \sup_{\|f\|_{\text{Lip}}\leqslant 1, \|f\|_{\text{BV}}\leqslant 1} \left\{ \left| \int f d\mu - \int f d\nu \right| \right\}. \qquad \texttt{def\_distanc}$$

Where the Lipschitz and the bounded variation norms are defined on the space of functions $f : \mathbb{T} \to \mathbb{R}$ as

$$\|f\|_{\text{Lip}} = \sup_{\substack{\theta_1,\theta_2\in\mathbb{T} \\ \theta_1\neq\theta_2}} \frac{|f\left(e^{i\theta_1}\right) - f\left(e^{i\theta_2}\right)|}{|e^{i\theta_1} - e^{i\theta_2}|},$$

$$\|f\|_{\text{BV}} = \sup_{n\geqslant 1, 0=\theta_1<\theta_2<\ldots<\theta_n=2\pi} \sum_{k=1}^{n-1} \left| f\left(e^{i\theta_{k+1}}\right) - f\left(e^{i\theta_k}\right) \right|.$$

The distance d is compatible with the weak convergence of probability measures [19]. We recall that for a $N \times N$ matrix A, its empirical measure of eigenvalues is defined by

$$\mu(A) = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j(A)},$$

where $\lambda_j(A)$, $j = 1, \ldots, N$, are the eigenvalues of A. The following Lemma, whose proof can be found in Appendix A, gives an upper bound on the distance of the empirical measures of two unitary matrices.

STANCE_INEQ **Lemma 2.2.** *For any A, B unitary matrices of size $N \times N$,*

- *For f with bounded variation,*

$$\left| \int f d\mu(A) - \int f d\mu(B) \right| \leqslant \|f\|_{BV} \frac{rank(A-B)}{N},$$

- *For f Lipschitz,*

$$\left| \int f d\mu(A) - \int f d\mu(B) \right| \leqslant \|f\|_{Lip} \frac{1}{N} \sum_{i,j=1}^{N} |(A-B)_{i,j}|.$$

As a consequence,

$$d(\mu(A),\mu(B)) \leqslant \min \left\{ \frac{rank(A-B)}{N}, \frac{1}{N} \sum_{i,j=1}^{N} |(A-B)_{i,j}| \right\}. \qquad \texttt{distance}$$

We are now in position to prove that the sequence $\mu_N(\mathcal{E})$ with $(\alpha_i)_{i\geqslant 1}$ *i.i.d* with law σ, such that $\text{supp}(\sigma) \subseteq \overline{\mathbb{D}}$, satisfies a large deviations principle. The proof of the following Lemma follows the same line as the corresponding one in [19].

ma weak ldp **Lemma 2.3.** *Let $(\alpha_i)_{i\geqslant 1}$ be an i.i.d sequence of law σ with supp$(\sigma) \subseteq \overline{\mathbb{D}}$, and $\mathcal{E}$ be the associated matrix defined in* (1.4)*. Then the sequence of empirical measures $(\mu_N(\mathcal{E}))_{N \text{ even}}$ satisfies a large deviations principle in $\mathcal{P}(\mathbb{T})$ endowed with the topology of weak convergence.*

Proof. We use a subadditivity argument to show that for any fixed $\mu \in \mathcal{P}(\mathbb{T})$ the following holds

$$\lim_{\delta\to 0} \liminf_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}\Big(\mu_N(\mathcal{E}) \in B_\mu(\delta)\Big) = \lim_{\delta\to 0} \limsup_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}\Big(\mu_N(\mathcal{E}) \in B_\mu(\delta)\Big), \tag{2.2}$$

eq weak ldp

where $B_\mu(\delta) := \{\nu \in \mathcal{P}(\mathbb{T}) \mid d(\mu,\nu) < \delta\}$. Then, applying [9, Theorem 4.1.11], along with the fact that in our setting a weak LDP is equivalent to a full LDP, due to the compactness of $\mathcal{P}(\mathbb{T})$, see remark 1.3, we conclude.
The first step to prove the result is to approximate the matrix $\mathcal{E}$ (whose law we denote by $\mathcal{E}^{(N)}$) by a diagonal block matrix of independent blocks. To this end, fix $q \in \mathbb{N}$ even such that $q \leqslant N$, write the euclidean division of N by q, $N = kq + r$ with $0 \leqslant r < q$. We consider $\mathcal{M}$ given by (1.3),

$$\mathcal{M} = \begin{pmatrix} -\alpha_N & & & & & \rho_N \\ & \Xi_2 & & & & \\ & & \Xi_4 & & & \\ & & & \ddots & & \\ & & & & \Xi_{N-2} & \\ \rho_N & & & & & \overline{\alpha}_N \end{pmatrix},$$

and approximate it the following way.

Let $\widetilde{\mathcal{M}} = \operatorname{diag}(\mathcal{M}_1, \dots, \mathcal{M}_k, R)$, where $\mathcal{M}_i$ is the block diagonal matrix given by

$$\mathcal{M}_i = \begin{pmatrix} -\widetilde{\alpha}_{(i-1)q} & & & & & \widetilde{\rho}_{(i-1)q} \\ & \Xi_{(i-1)q+2} & & & & \\ & & \Xi_{(i-1)q+4} & & & \\ & & & \ddots & & \\ & & & & \Xi_{iq-2} & \\ \widetilde{\rho}_{(i-1)q} & & & & & \overline{\widetilde{\alpha}}_{(i-1)q} \end{pmatrix},$$

where $(\widetilde{\alpha}_{(i-1)q})_{1\leqslant i\leqslant k}$ are *i.i.d* of law σ, independent of the α_i's, $\widetilde{\rho}_i = \sqrt{1-|\widetilde{\alpha}_i|^2}$, and the remaining block (of size $r \times r$) R is defined similarly:

$$R = \begin{pmatrix} -\widetilde{\alpha}_{kq} & & & & & \widetilde{\rho}_k \\ & \Xi_{kq+2} & & & & \\ & & \Xi_{kq+4} & & & \\ & & & \ddots & & \\ & & & & \Xi_{N-2} & \\ \widetilde{\rho}_{kq+1} & & & & & \overline{\widetilde{\alpha}_{kq}} \end{pmatrix}.$$

Following the same decomposition of $N = kq + r$ we write $\mathcal{L} = \operatorname{diag}(\mathcal{L}_1, \dots, \mathcal{L}_k, \mathcal{L}_{k+1})$, with $\mathcal{L}_i$ of size q for $1 \leqslant i \leqslant k$ and $\mathcal{L}_{k+1}$ of size r.
Notice that by construction, we have

$$\operatorname{rank}(\mathcal{M} - \widetilde{\mathcal{M}}) \leqslant 2(k+1). \tag{2.3}$$

ineq_rank_t

Now, defining $\widetilde{\mathcal{E}} = \mathcal{L}\widetilde{\mathcal{M}}$, $\widetilde{\mathcal{E}}$ is a block diagonal matrix $\operatorname{diag}(\mathcal{E}_1, \dots, \mathcal{E}_k, \mathcal{E}_{k+1})$.Then, the blocks $\mathcal{E}_i$, $1 \leqslant i \leqslant k+1$ are independent, each $\mathcal{E}_i$, $i = i, \dots, k$, has law $\mathcal{E}^{(q)}$, and $\mathcal{E}_{k+1}$ has law $\mathcal{E}^{(r)}$.
Furthermore, using that $\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank}(A); \operatorname{rank}(B)\}$ for A, B two square matrices, and (2.3) we get

$$\operatorname{rank}(\mathcal{E} - \tilde{\mathcal{E}}) = \operatorname{rank}(\mathcal{L}(\mathcal{M} - \widetilde{\mathcal{M}})) \leqslant 4(k+1).$$

By the first point of Lemma 2.2 we deduce

$$d(\mu_N(\mathcal{E}), \mu_N(\widetilde{\mathcal{E}})) \leqslant \frac{4(k+1)}{N} \leqslant \frac{8}{q}.$$

Moreover, we can rewrite $\mu_N(\widetilde{\mathcal{E}})$ as

$$\mu_N(\widetilde{\mathcal{E}}) = \frac{q}{N}\sum_{\ell=1}^{k}\mu_q(\mathcal{E}_\ell) + \frac{r}{N}\mu_r(\mathcal{E}_{k+1})\,.$$

Using the independence of the blocks of $\widetilde{\mathcal{E}}$, we deduce that

$$\begin{aligned}\mathbb{P}\Big(\mu_q(\mathcal{E}_1)\in B_\mu(\delta)\Big)^k\mathbb{P}\Big(\mu_r(\mathcal{E}_{k+1})\in B_\mu(\delta)\Big) &= \mathbb{P}\Big(\mu_q(\mathcal{E}_1),\dots,\mu_q(\mathcal{E}_k),\mu_r(\mathcal{E}_{k+1})\in B_\mu(\delta)\Big)\\ &\leqslant \mathbb{P}\Big(\frac{q}{N}\sum_{l=1}^{k}\mu_q(\mathcal{E}_l)+\frac{r}{N}\mu_r(\mathcal{E}_{k+1})\in B_\mu(\delta)\Big)\\ &= \mathbb{P}\Big(\mu_N(\tilde{\mathcal{E}})\in B_\mu(\delta)\Big)\\ &\leqslant \mathbb{P}\Big(\mu_N(\mathcal{E})\in B_\mu\Big(\delta+\frac{8}{q}\Big)\Big),\end{aligned}$$

Where we used the convexity of balls in the first inequality.
This implies that, setting

$$u_N(\delta) = -\ln\left(\mathbb{P}(\mu_N\in B_\mu(\delta))\right)\,,$$

we have

$$u_N\left(\delta+\frac{8}{q}\right)\leqslant k u_q(\delta)+u_r(\delta).$$

We now conclude as in [19, Lemma 2.3]. Let $\delta>0$ and choose q in such a way that $\frac{8}{q}\leqslant\delta$, so we deduce that

$$\frac{u_N(2\delta)}{N}\leqslant\frac{u_N\left(\delta+\frac{8}{q}\right)}{N}\leqslant\frac{u_q(\delta)}{q}+\frac{u_r(\delta)}{N}\,,$$

since $\frac{u_r(\delta)}{N}\to 0$ as $N\to\infty$, we deduce that

$$\limsup_{N\to\infty}\frac{u_N(2\delta)}{N}\leqslant\frac{u_q(\delta)}{q}\,.$$

The previous inequality holds true for all q big enough, so we conclude that

$$\limsup_{N\to\infty}\frac{u_N(2\delta)}{N}\leqslant\liminf_{N\to\infty}\frac{u_N(\delta)}{N}\,.$$

From this last inequality we deduce that

$$\lim_{\delta\to 0}\limsup_{N\to\infty}\frac{u_N(\delta)}{N}\leqslant\lim_{\delta\to 0}\liminf_{N\to\infty}\frac{u_N(\delta)}{N}\,,$$

thus we obtain (2.2), and the conclusion follows applying [9, Theorem 4.1.11]. □

Since $(\mathcal{P}(\mathbb{T}),d)$ is compact, Lemma 2.3 automatically implies the existence of a **strong** large deviations principle. Furthermore, the corresponding rate function J ,which depends on the distribution σ of the entries of $\mathcal{L}$ and $\mathcal{M}$, can be seen to be convex. We collect these results in the following proposition.

deviations

Proposition 2.4. *In the same hypothesis and notations as in Lemma 2.3, the sequence of empirical measures $(\mu_N(\mathcal{E}))_{N\ even}$ satisfies a large deviations principle with good, convex rate function $J:\mathcal{P}(\mathbb{T})\to[0,+\infty]$, i.e.*

- *The function J is convex and its level sets $J^{-1}([0,a])$, $a\geqslant 0$, are compact,*

- *For all $O \subset \mathcal{P}(\mathbb{T})$ open,*

$$-\inf_O J \leqslant \liminf_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}(\mu_N(\mathcal{E}) \in O),$$

- *For all $F \subset \mathcal{P}(\mathbb{T})$ closed,*

$$\limsup_{N \text{ even}} \frac{1}{N} \ln \mathbb{P}(\mu_N(\mathcal{E}) \in F) \leqslant -\inf_F J.$$

Proof. We already established all the claims except the fact that the function J is convex and that the level sets $J^{-1}([0,a])$ are compact. The latter comes from the fact that these sets are closed, see [9, Theorem 4.1.11]. To prove the convexity of J, we follow the same argument as [19, Theorem 2.4].

Let $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{T})$. Since $\mu_{2N}(\mathcal{E})$ can be approximated by the sum of two independent $\mu_N(\mathcal{E})$'s up to a mistake smaller than $\frac{4}{N}$ by the first point of Lemma 2.2, for $\delta > 0$ the following holds

$$\mathbb{P}\left(\mu_N(\mathcal{E}) \in B_{\mu_1}(\delta)\right) \mathbb{P}\left(\mu_N(\mathcal{E}) \in B_{\mu_2}(\delta)\right) \leqslant \mathbb{P}\left(\mu_{2N}(\mathcal{E}) \in B_{\frac{\mu_1+\mu_2}{2}}\left(\delta + \frac{4}{N}\right)\right),$$

taking minus the logarithm of both sides, dividing by $2N$, taking the limit for N going to infinity and then for δ to zero, we deduce that:

$$J\left(\frac{\mu_1+\mu_2}{2}\right) \leqslant \frac{1}{2}\left(J(\mu_1) + J(\mu_2)\right),$$

which, together with the lower semi-continuity of J, implies the convexity of J, see [9, Lemma 4.1.21] . □

2.2 Large deviations principle for the Ablowitz-Ladik lattice

Taking $\sigma = \Theta_{2\beta+1}$ given by equation (2.1), Proposition 2.4 applies to $(\mu_N(\mathcal{E}))_{N \text{ even}}$, where $\mathcal{E}$ follows $\mathbb{P}^{\beta}_{AL,N}$ defined in (1.6). Thus, $(\mu_N(\mathcal{E}))_{N \text{ even}}$ with law $\mathbb{P}^{\beta}_{AL,N}$ satisfies a large deviations principle, with a good convex rate function, that we denote by J_β.

We can now state the existence of a large deviations principle for $(\mu_N(\mathcal{E}))_{N \text{ even}}$ under $\mathbb{P}^{V,\beta}_{AL,N}$ for V continuous.

cor: LDP AL **Corollary 2.5.** *Let $\beta > 0$, and $V : \mathbb{T} \to \mathbb{R}$ be continuous. Under $\mathbb{P}^{V,\beta}_{AL,N}$ the sequence $(\mu_N(\mathcal{E}))_{N \text{ even}}$ fulfils a large deviations principle with good, convex rate function $J^V_\beta(\mu) = g^V_\beta(\mu) - \inf_{\nu \in \mathcal{P}(\mathbb{T})} g^V_\beta(\nu)$, where $g^V_\beta(\mu)$ is given for $\mu \in \mathcal{P}(\mathbb{T})$ by*

$$g^V_\beta(\mu) = J_\beta(\mu) + \int_{\mathbb{T}} V \mathrm{d}\mu\,. \tag{2.4}$$ eq: rate f

Proof. Let us write

$$\mathrm{d}\mathbb{P}^{V,\beta}_{AL,N} = \frac{Z^{AL}_N(0,\beta)}{Z^{AL}_N(V,\beta)} e^{-N\int_{\mathbb{T}} V \mathrm{d}\mu_N} \mathrm{d}\mathbb{P}^{\beta}_{AL,N} = \frac{1}{\mathcal{Z}^{AL,V}_N} e^{-N\int_{\mathbb{T}} V \mathrm{d}\mu_N} \mathrm{d}\mathbb{P}^{\beta}_{AL,N}.$$

The function $\mu \mapsto \int_{\mathbb{T}} V \mathrm{d}\mu$ being bounded continuous, by the large deviations principle under $\mathbb{P}^{\beta}_{AL,N}$ and Varadhan's Lemma, [13, Theorem 1.2.1], we see that for any bounded continuous $f : \mathcal{P}(\mathbb{T}) \to \mathbb{R}$ we have

$$\lim_N \frac{1}{N} \ln \int_{\mathbb{D}^N} e^{Nf(\mu_N)} \mathrm{d}\mathbb{P}^{V,\beta}_{AL,N} = \sup_{\mu \in \mathcal{P}(\mathbb{T})} \left\{ f(\mu) - \left(J_\beta(\mu) + \int V \mathrm{d}\mu - \inf_{\nu \in \mathcal{P}(\mathbb{T})} \left\{ J_\beta(\nu) + \int V \mathrm{d}\nu \right\} \right) \right\},$$

which ensures by [13, Theorem 1.2.3] that (μ_N) satisfies a large deviations principle under $\mathbb{P}^{V,\beta}_{AL,N}$ with the announced rate function. Since the function J^V_β is an affine perturbation of J_β, which is convex, J^V_β is also convex. □

The first point of Theorem 1.4 is proven.

3 Circular β ensemble at high temperature

sec Circular

In this section, we consider the Circular β ensemble, and we collect some known results that we exploit in our treatment. The aim of this section is to prove an alternative formulation of the large deviations principle for the Circular beta ensemble in the high-temperature regime, see Theorem 3.11 below. Our formulation allows us to relate the large deviations principle of the Coulomb gas with the one of Ablowitz-Ladik, proved in the previous section.

3.1 Large deviations principle for Circular β ensemble

Coulomb gas on the torus $\mathbb{T} = [-\pi, \pi)$ at temperature $\widetilde{\beta}^{-1}$ are described by

$$\mathrm{d}\mathbb{P}^{V,\widetilde{\beta}}_{C,N} = \frac{1}{Z^C_N(V,\widetilde{\beta})} \prod_{j\leqslant\ell\leqslant N} |e^{i\theta_j} - e^{i\theta_\ell}|^{\widetilde{\beta}} e^{\sum_{j=1}^N V(\theta_j)} \mathrm{d}\boldsymbol{\theta}\,, \tag{3.1}$$

eq:circular

here $V : \mathbb{T} \to \mathbb{R}$ is a continuous potential, and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$. When $V = 0$, Killip and Nenciu showed that $\mathrm{d}\mathbb{P}^{0,\widetilde{\beta}}_{C,N}$ is the law of the eigenvalues of a CMV matrix [21], see Theorem 3.4. In this manuscript, we are interested in the so-called *high-temperature regime* of this ensemble, namely the limit of number of particles N going to infinity with the constraint that $\widetilde{\beta}N \to 2\beta > 0$. This regime was considered by Hardy and Lambert in [20], who proved the following large deviations principle for the measure $\mu_N = \frac{1}{N}\sum_{j=1}^N \delta_{e^{i\theta_j}}$, where the θ_j are distributed according to (3.1).

n_LDPCoulomb

Theorem 3.1. *Let $\widetilde{\beta} = \frac{2\beta}{N}$, $\beta > 0$ and assume $V : \mathbb{T} \to \mathbb{R}$ to be continuous. Define for any $\mu \in \mathcal{P}(\mathbb{T})$ absolutely continuous with respect to the Lebesgue measure the functional*

$$f^V_\beta(\mu) = -\beta \int_{\mathbb{T}\times\mathbb{T}} \log\left(|e^{i\theta} - e^{i\varphi}|\right) \mu(\mathrm{d}\theta)\mu(\mathrm{d}\varphi) + \beta\log(2) + \int_{\mathbb{T}} V(\theta)\mu(\mathrm{d}\theta) + \int_{\mathbb{T}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta)\right)\mu(\mathrm{d}\theta) + \log(2\pi)\,,$$

then

i. *the functional $f^V_\beta(\mu)$ is strictly convex and achieves its minimal value at the unique probability measure μ^V_β absolutely continuous with respect to the Lebesgue measure;*

ii. *the sequence (μ_N) satisfies a large deviations principle in $\mathcal{P}(\mathbb{T})$ equipped with the weak topology at speed βN with rate function defined for absolutely continuous $\mu \in \mathcal{P}(\mathbb{T})$ with respect to Lebesgue measure by $I^V_\beta(\mu) = f^V_\beta(\mu) - f^V_\beta(\mu^V_\beta)$, and $I^V_\beta(\mu) = +\infty$ otherwise. In particular*

$$\mu_N \xrightarrow[N\to\infty]{a.s.} \mu^V_\beta\,.$$

Exploiting this result, in [24] the authors deduced several useful properties of the minimizer μ^V_β, specifically they proved the following.

prop circ

Lemma 3.2 (cf. [24] Lemma 3.5)**.** *Let $\beta > 0$, consider a continuous potential $V : \mathbb{T} \to \mathbb{R}$, then the following holds*

i. *The map $\beta \to \inf\left(f^V_\beta(\mu)\right)$ is Lipschitz;*

ii. Let D be the distance on $\mathcal{P}(\mathbb{T})$ given by

$$D(\mu,\mu') = \left(-\int_{\mathbb{T}\times\mathbb{T}} \ln\left|\sin\left(\frac{\theta-\phi}{2}\right)\right| (\mu-\mu')(\mathrm{d}\theta)(\mu-\mu')(\mathrm{d}\phi)\right)^{1/2} = \sqrt{\sum_{k\geqslant 1}\frac{1}{k}\left|\widehat{\mu}_k-\widehat{\mu}'_k\right|^2}, \tag{eq:distance}$$

where $\widehat{\mu}_k = \int_{\mathbb{T}} e^{ik\theta}\mu(\mathrm{d}\theta)$. Then for any $\varepsilon>0$ there exists a finite constant C_ε such that for all $\beta,\beta'>\varepsilon$

$$D(\mu_\beta^V,\mu_{\beta'}^V) \leqslant C_\varepsilon\left|\beta-\beta'\right|. \tag{3.2}$$ Distance

lip_moments **Remark 3.3.** *We observe that if $f\in L^2(\mathbb{T})$ with derivative in $L^2(\mathbb{T})$, we can set $||f||_{\frac{1}{2}} = \sqrt{\sum_{k\geqslant 1} k|\widehat{f}_k|^2}$. So, for any measure ν with zero mass we obtain the following*

$$\int_{\mathbb{T}} f(\theta)\nu(\mathrm{d}\theta) = \sum_{k\neq 0}\widehat{f}_k\overline{\widehat{\nu}_k} = \sum_{k\neq 0} k\widehat{f}_k\frac{\overline{\widehat{\nu}_k}}{k}.$$

Then, by Cauchy-Schwartz inequality, we deduce the following inequality

$$\left|\int_{\mathbb{T}} f(\theta)\nu(\mathrm{d}\theta)\right|^2 \leqslant \left|\sum_{k\neq 0} k\widehat{f}_k\sum_{k\neq 0}\frac{\overline{\widehat{\nu}_k}}{k}\right| \leqslant 4||f||_{\frac{1}{2}}^2 D(\nu,0)^2. \tag{3.3}$$ eq:CS

Combining (3.2) *and* (3.3)*, we deduce that for any function f with finite $||f||_{\frac{1}{2}}$ norm, the map $\beta\to\int_{\mathbb{T}} f\mathrm{d}\mu_\beta^V(\theta)$ is Lipschitz for $\beta>0$.*

3.2 Relation with the large deviations principle of the Ablowitz-Ladik lattice

In the case $V=0$, for any $\widetilde{\beta}>0$, Killip and Nenciu in [21] showed that the law $\mathbb{P}_{C,N}^{0,\widetilde{\beta}}$ (3.1) coincides with the distribution of the eigenvalues of a certain CMV matrix. Specifically they proved the following:

Killipenciu **Theorem 3.4** (cf. [21] Theorem 1.2)**.** *Consider the block diagonal $N\times N$ matrices*

$$L = diag\,(\Xi_1,\Xi_3,\Xi_5\ldots,) \quad \text{and} \quad M = diag\,(\Xi_0,\Xi_2,\Xi_4,\ldots), \tag{eq:LME}$$

where the block Ξ_j, $j=1,\ldots,N-1$, takes the form

$$\Xi_j = \begin{pmatrix}\overline{\alpha}_j & \rho_j\\ \rho_j & -\alpha_j\end{pmatrix}, \quad \rho_j=\sqrt{1-|\alpha_j|^2}, \tag{eq:xi_def}$$

while $\Xi_0=(1)$ and $\Xi_N=(\overline{\alpha}_N)$ are 1×1 matrices. Define the $N\times N$ sparse matrix

$$E = LM, \tag{3.4}$$ E

and suppose that the entries α_j are independent complex random variables with $\alpha_j\sim\Theta_{\widetilde{\beta}(N-j)+1}$ for $1\leqslant j\leqslant N-1$ and α_N is uniformly distributed on the unit circle. Then the eigenvalues of E are distributed according to the Circular Ensemble (3.1) *at temperature $\widetilde{\beta}^{-1}$.*

To simplify the notation, we will denote by $\mathbb{P}_{C,N}^{\widetilde{\beta}}$ the law $\mathbb{P}_{C,N}^{0,\widetilde{\beta}}$. We give an alternative formulation of the large deviations principle for the empirical measure under the law $\mathbb{P}_{C,N}^{\frac{2\beta}{N}}$ based on the Killip-Nenciu matrix representation. This alternative formulation allows us to relate the

rate function of the Coulomb gas I_β in terms of the rate function J_β of the Ablowitz-Ladik lattice. Finally, applying Varadhan's Lemma [13, Theorem 1.2.1] we obtain an alternative formulation of the large deviations principle for the Circular beta ensemble at high temperature with continuous potential, see Theorem 3.11 below.

To achive our goal, we need several technical results regarding the distribution Θ_ν (2.1), and the CMV matrix E (3.4). First, in the next Lemma, we give a reprensentation of Θ_ν in terms of Gaussian, and Chi distributions.

resentation **Lemma 3.5.** *Let $\nu > 1$. Let X_1, X_2, Y_ν be independent, X_1, X_2 standard Gaussian variables and Y_ν be $\chi_{\nu-1}$ distributed, i.e. with density*

$$\chi_{\nu-1}(x) = \frac{2^{\frac{3-\nu}{2}}}{\Gamma(\frac{\nu-1}{2})} x^{\nu-2} e^{-x^2/2} \mathbf{1}_{x>0}\,,$$

here $\Gamma(x)$ is the classical Gamma function [10, §5]

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt\,, \tag{3.5}$$

eq gamma

Then, $Z := \frac{X_1+iX_2}{(X_1^2+X_2^2+Y_\nu^2)^{1/2}}$ follows the law Θ_ν.

Remark 3.6. *If $\nu \geqslant 2$ is an integer, Y_ν^2 has the law of $\sum_{i=1}^{\nu-1} N_i^2$ where the N_i's are i.i.d. standard gaussians random variables, thus Z is equal in distribution to $\frac{X_1+iX_2}{(X_1^2+\cdots+X_{\nu+1}^2)^{\frac{1}{2}}}$, which follows the law Θ_ν by Remark 2.1.*

Proof. We identify $\mathbb{C}$ with $\mathbb{R}^2$ and check that for any $f : \mathbb{D} \to \mathbb{R}$ bounded and measurable,

$$\mathbb{E}[f(Z)] = \frac{\nu-1}{2\pi} \int_{\mathbb{D}} f(z)(1-|z|^2)^{\frac{\nu-3}{2}} \mathrm{d}^2 z,$$

i.e. that for some constant c,

$$\int_{\mathbb{R}^2\times\mathbb{R}_+^*} f\left(\frac{x_1}{(x_1^2+x_2^2+y^2)^{1/2}}, \frac{x_2}{(x_1^2+x_2^2+y^2)^{1/2}}\right) e^{-\frac{x_1^2+x_2^2}{2}} e^{-\frac{y^2}{2}} y^{\nu-2} \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}y$$
$$= c \int_{\mathbb{D}} f(u,v)(1-(u^2+v^2))^{\frac{\nu-3}{2}} \mathrm{d}u\mathrm{d}v\,.$$

For fixed $y > 0$, we perform the diffeomorphic change of variables $(u,v) = \frac{1}{(x_1^2+x_2^2+y^2)^{1/2}}(x_1,x_2)$. Its inverse $(x_1,x_2) = \frac{y}{(1-(u^2+v^2))^{1/2}}(u,v)$ has Jacobian equal to $y^2(1-(u^2+v^2))^{-2}$. The integral becomes

$$\int_{\mathbb{D}} \frac{f(u,v)}{(1-(u^2+v^2))^2} \int_{\mathbb{R}_+^*} e^{-\frac{y^2}{2(1-(u^2+v^2))}} y^\nu \mathrm{d}y\mathrm{d}u\mathrm{d}v = \sqrt{\frac{\pi}{2}} \int_{\mathbb{D}} \frac{f(u,v)}{(1-(u^2+v^2))^{3/2}} \mathbb{E}[|X_{u,v}|^\nu] \mathrm{d}u\mathrm{d}v\,,$$

where $X_{u,v}$ denotes a Gaussian variable $\mathcal{N}(0, 1-(u^2+v^2))$. By [34],

$$\mathbb{E}[|X_{u,v}|^\nu] = c_\nu (1-(u^2+v^2))^{\frac{\nu}{2}}$$

for some constant c_ν independent of u, v. Substituting this last equality in the previous integral, we conclude.

□

To obtain our main result, we need some technical lemmas. Since they are based on standard techniques, we just state them here, and we defer their proofs to the appendix A.

The first one gives an estimate that we use combined with Lemma 2.2.

ate_product **Lemma 3.7.** *Let $N = 2k$ be even and A be a $N \times N$ matrix. Then,*

- $\sum_{i,j=1}^{N} |(\mathcal{L}A)_{i,j}| \leqslant 2\sum_{i,j=1}^{N} |A_{i,j}|$,
- $\sum_{i,j=1}^{N} |(A\mathcal{M})_{i,j}| \leqslant 2\sum_{i,j=1}^{N} |A_{i,j}|$,

where $\mathcal{M}$, and $\mathcal{L}$ are defined in (1.3).

We now give an explicit coupling between Θ_ν (2.1) and $\Theta_{\nu+h}$ for $\nu > 1$, $h > 0$.
Let X_1, X_2 be $\mathcal{N}(0,1)$ independent variables, and let $Y_{\nu-1} \sim \chi_{\nu-1}$, $Y_h \sim \chi_h$ be independent, and independent of X_1, X_2 (notice that $(Y_h^2 + Y_{\nu-1}^2)^{\frac{1}{2}}$ is $\chi_{\nu+h-1}$ distributed). Let

$$\alpha_\nu = \frac{X_1 + iX_2}{(X_1^2 + X_2^2 + Y_{\nu-1}^2)^{\frac{1}{2}}}, \qquad \alpha_{\nu+h} = \frac{X_1 + iX_2}{(X_1^2 + X_2^2 + Y_{\nu-1}^2 + Y_h^2)^{\frac{1}{2}}}. \tag{3.6}$$

eq_coupling

By Lemma 3.5, $\alpha_\nu \sim \Theta_\nu$ and $\alpha_{\nu+h} \sim \Theta_{\nu+h}$.

Exploiting this coupling, we bound the differences $|\alpha_\nu - \alpha_{\nu+h}|$ and $|\rho_\nu - \rho_{\nu+h}|$ by a random variable Z_h, where $\rho_\nu = \sqrt{1 - |\alpha_\nu|^2}$, and $\rho_{\nu+h} = \sqrt{1 - |\alpha_{\nu+h}|^2}$. Moreover, we find an upper bound for the exponential moments of Z_h.

lem:bomb **Lemma 3.8.** *Let α_ν and $\alpha_{\nu+h}$ defined by equation* (3.6). *Define $\rho_\nu = \sqrt{1 - |\alpha_\nu|^2}$, and $\rho_{\nu+h} = \sqrt{1 - |\alpha_{\nu+h}|^2}$, then the following holds*

i.

$$|\alpha_\nu - \alpha_{\nu+h}| \leqslant \frac{Y_h}{(X_1^2 + X_2^2 + Y_h^2)^{\frac{1}{2}}} \quad \textit{almost surely},$$
$$|\rho_\nu - \rho_{\nu+h}| \leqslant \frac{Y_h}{(X_1^2 + X_2^2 + Y_h^2)^{\frac{1}{2}}} \quad \textit{almost surely},$$

where $X_1, X_2 \sim \mathcal{N}(0,1)$, $Y_h \sim \chi_h$ are all independent.

ii. define $Z_h = \frac{Y_h}{(X_1^2+X_2^2+Y_h^2)^{\frac{1}{2}}}$, and $a(h) = -\frac{1}{2}\ln(h) + 1$, then there exists a constant K independent of h such that

$$\sup_{0<h<1} \mathbf{E}\left[\exp(a(h)Z_h)\right] \leqslant K\,. \tag{3.7}$$

eq:sup_boun

ne coupling **Remark 3.9.** *Let $h < h'$, and let $Z_h, Z_{h'}$ be given by*

$$Z_h = \frac{Y_h}{(X_1^2 + X_2^2 + Y_h^2)^{\frac{1}{2}}}, \qquad Z_{h'} = \frac{Y_{h'}}{(X_1^2 + X_2^2 + Y_{h'}^2)^{\frac{1}{2}}}\,,$$

where $Y_h \sim \chi_h$ and $Y_{h'} \sim \chi_{h'}$ are χ variables coupled by

$$Y_{h'} = \sqrt{Y_h^2 + Z^2},$$

Z being a $\chi_{h'-h}$ variable independent of Y_h. Then, because of the monotonicity of the function $x \mapsto \frac{x}{\sqrt{a+x^2}}$ for $a > 0$, we have almost surely $Z_h \leqslant Z_{h'}$.

We are now in position to give an alternate formulation of the large deviations principle for the sequence of measures $(\mu_N(E))$ under the law $\mathbb{P}_{C,N}^{\frac{2\beta}{N}}$, given by Theorem 3.1.

Lemma 3.10. *Let $\beta > 0$. The law of the empirical measure $(\mu_N(E))_{N\,even}$ under $\mathbb{P}_{C,N}^{\frac{2\beta}{N}}$ satisfies a large deviations principle at speed N and with a good rate function*

$$I_\beta(\mu) = \lim_{\delta\to 0}\liminf_{q\to\infty} \inf_{\substack{\nu_{\beta/q},\dots,\nu_\beta \\ \frac{1}{q}\sum_i \nu_{i\beta/q}\in B_\mu(\delta)}} \left\{ \frac{1}{q}\sum_{i=1}^{q} J_{i\beta/q}(\nu_{i\beta/q}) \right\},$$

where J_η is the rate function of Proposition 2.4 applied to $\sigma = \Theta_{2\eta+1}$.

Proof. Following the same line as in [19, Lemma 3.3], we proceed by exponential approximation. Let $q \geqslant 1$ be an integer, since N is even, we can write $N = kq + r$, with k even, and with $0 \leqslant r < 2q-2$. Consider the following family of matrices $\mathcal{L}^{(i)}, \mathcal{M}^{(i)}$, $i = 1, \dots q$ defined as

$$\mathcal{M}_k^{(i)} = \begin{pmatrix} -\alpha_{ik} & & & & & \rho_{ik} \\ & \Xi_2^{(i)} & & & & \\ & & \Xi_4^{(i)} & & & \\ & & & \ddots & & \\ & & & & \Xi_{k-2}^{(i)} & \\ \rho_{ik} & & & & & \overline{\alpha}_{ik} \end{pmatrix}, \qquad \mathcal{L}_k^{(i)} = \begin{pmatrix} \Xi_1^{(i)} & & & \\ & \Xi_3^{(i)} & & \\ & & \ddots & \\ & & & \Xi_{k-1}^{(i)} \end{pmatrix},$$

where $\Xi_\ell^{(i)}$ are defined as

$$\Xi_\ell^{(i)}(\alpha_{i(k-1)+\ell}) = \Xi_\ell^{(i)} = \begin{pmatrix} \overline{\alpha}_{i(k-1)+\ell} & \rho_{i(k-1)+\ell} \\ \rho_{i(k-1)+\ell} & -\alpha_{i(k-1)+\ell} \end{pmatrix}, \quad \rho_{i(k-1)+\ell} = \sqrt{1-|\alpha_{i(k-1)+\ell}|^2}, \quad \ell = 1, \dots, k-1\,, \tag{3.8}$$

eq: local

and $(\alpha_{i(k-1)+\ell})_{1\leqslant i\leqslant q, 1\leqslant \ell\leqslant k}$ is a family of independent random variables such that

$$\alpha_{i(k-1)+\ell} \sim \Theta_{2\beta\frac{N-ik}{N}+1},\ \ell = 1, \dots k,\ i = 1, \dots, q.$$

From these two families of matrices, we can define a third one, namely $\mathcal{E}_k^{(i)} = \mathcal{L}_k^{(i)}\mathcal{M}_k^{(i)}$, $i = 1, \dots, q$. We notice that $\mathcal{E}_k^{(i)}$ is distributed according to $\mathbb{P}_{AL,k}^{\beta\frac{N-ik}{N}}$, and that the $\mathcal{E}_k^{(i)}$, $1 \leqslant i \leqslant q$, are independent.

Our aim is to prove that the empirical measure of the matrix C_N^q

$$C_N^q = \begin{pmatrix} \mathcal{E}_k^{(1)} & & & & \\ & \mathcal{E}_k^{(2)} & & & \\ & & \ddots & & \\ & & & \mathcal{E}_k^{(q)} & \\ & & & & 0_{r\times r} \end{pmatrix},$$

where $0_{r\times r}$ is a null block of size $r \times r$, is an exponential approximation (see [9, Definition 4.2.14]) of the empirical measure of $E \sim \mathbb{P}_{C;N}^{\frac{2\beta}{N}}$ (3.4), that is, for any positive real number δ:

$$\lim_{q\to\infty}\limsup_{N\to\infty} \frac{1}{N}\ln\left(\mathbb{P}\left(d\left(\mu_N(E), \mu_N(C_N^q)\right)\right) > \delta\right) = -\infty\,, \tag{3.9}$$

eq:real cla

where $\mathbb{P}$ denotes the coupling introduced in equation (3.6). In this way, we obtain the claim as an application of [9, Theorem 4.2.16]. The strategy of proof is the following. First we approximate C_N^q and E by two block diagonal matrices $\widetilde{C}_N^q, \widetilde{E}_N^q$ respectively. Finally, we will prove that both $\widetilde{C}_N^q$, and $\widetilde{E}_N^q$ approximate a third matrix B_N^q.

Consider another family of matrices $(\widetilde{\mathcal{M}}_k^{(i)})_{1\leqslant i\leqslant q}$ of size $k\times k$, defined as

$$\widetilde{\mathcal{M}}_k^{(i)} = \operatorname{diag}\left(\widetilde{\Xi}_0^{(i)}, \Xi_2^{(i)}, \Xi_4^{(i)}, \dots, \widetilde{\Xi}_k^{(i)}\right),$$

where the matrices $\Xi_\ell^{(i)}$ are defined in (3.8), while $\widetilde{\Xi}_0^{(i)} = (1)$ and $\widetilde{\Xi}_k^{(i)} = (\overline{\alpha}_{ik})$ are 1×1 matrices, where the α_{ik} are independent, uniformly distributed on the unit circle for all $i = 1,\dots,q$, and independent from $(\alpha_{i(k-1)+l})_{1\leqslant i\leqslant q, 1\leqslant \ell\leqslant k-1}$. Define the $k\times k$ family of CMV matrices

$$\widetilde{\mathcal{E}}_k^{(i)} = \mathcal{L}_k^{(i)}\widetilde{\mathcal{M}}_k^{(i)} \quad i = 1,\dots,q\,.$$

From the family of matrices $(\widetilde{\mathcal{E}}_k^{(i)})_{1\leqslant i\leqslant q}$, we define the block diagonal matrix:

$$\widetilde{C}_N^q = \begin{pmatrix} \widetilde{\mathcal{E}}_k^{(1)} & & & & \\ & \widetilde{\mathcal{E}}_k^{(2)} & & & \\ & & \ddots & & \\ & & & \widetilde{\mathcal{E}}_k^{(q)} & \\ & & & & 0_{r\times r} \end{pmatrix}.$$

We claim that $\widetilde{C}_N^q$ is such that

$$\operatorname{rank}(C_N^q - \widetilde{C}_N^q) \leqslant 2q\,. \tag{3.9}$$

eq:rank_1

Indeed, we take the same $\alpha_j^{(i)}$ in the construction of $\widetilde{\mathcal{E}}_k^i$ and of $\mathcal{E}_k^i$, except for the entries of the corners of $\mathcal{M}_k^{(i)}$, where $\mathcal{M}_k^{(i)}(1,1)$ is replaced by 1, $\mathcal{M}_k^{(i)}(k,k)$ is replaced by a uniform variable on the circle, and both entries $\mathcal{M}_k^{(i)}(1,k)$ and $\mathcal{M}_k^{(i)}(k,1)$ are replaced by 0. This shows that

$$\operatorname{rank}(\mathcal{M}_k^{(i)} - \widetilde{\mathcal{M}}_k^{(i)}) \leqslant 2,$$

and

$$\operatorname{rank}(\mathcal{E}_k^{(i)} - \widetilde{\mathcal{E}}_k^{(i)}) = \operatorname{rank}(\mathcal{L}_k^{(i)}(\mathcal{M}_k^{(i)} - \widetilde{\mathcal{M}}_k^{(i)})) \leqslant \operatorname{rank}(\mathcal{M}_k^{(i)} - \widetilde{\mathcal{M}}_k^{(i)}) \leqslant 2,$$

and we deduce (3.9). From (3.9) and Lemma 2.2, we deduce that

$$d(\mu_N(C_N^q), \mu_N(\widetilde{C}_N^q)) \leqslant \frac{2}{k}\,,$$

and for any $\delta > 0$ and sufficiently large N, we can take k such that $\frac{2}{k} \leqslant \frac{\delta}{4}$.

Consider now another two families of matrices $(\mathfrak{L}_k^{(i)})_{1\leqslant i\leqslant q}$, and $(\mathfrak{M}_k^{(i)})_{1\leqslant i\leqslant q}$, constructed in the same way as $(\mathcal{L}_k^{(i)})_{1\leqslant i\leqslant q}$, and $(\widetilde{\mathcal{M}}_k^{(i)})_{1\leqslant i\leqslant q}$ by means of independent variables $\widetilde{\alpha}_{(i-1)k+j}$, where each $\widetilde{\alpha}_{(i-1)k+j} \sim \Theta_{2\beta\frac{N-(i-1)k-j}{N}}$ is coupled to $\alpha_{(i-1)k+j}$ by equation (3.6), for all $j = 0,\dots,k-1$, and $i = 1,\dots,q$, and where $\widetilde{\alpha}_{ik} = \alpha_{ik}$ for $i = 1,\dots,q$ is uniformly distributed on the unit circle. Define the family of CMV matrices $(\mathfrak{E}_k^{(i)})_{1\leqslant i\leqslant q}$ as

$$\mathfrak{E}_k^{(i)} = \mathfrak{L}_k^{(i)}\mathfrak{M}_k^{(i)}\,, \quad i = 1,\dots,q\,.$$

Define the block diagonal matrix E_N^q as:

$$E_N^q = \begin{pmatrix} \mathfrak{E}_k^{(1)} & & & & \\ & \mathfrak{E}_k^{(2)} & & & \\ & & \ddots & & \\ & & & \mathfrak{E}_k^{(q)} & \\ & & & & 0_{r\times r}\,. \end{pmatrix}$$

From the definition of $\mathbb{P}_{C,N}^{\frac{2\beta}{N}}$ and E_N^q, we conclude that for some $E \sim \mathbb{P}_{C,N}^{\frac{2\beta}{N}}$, we have

$$\operatorname{rank}(E - E_N^q) \leqslant 2q + r\,.$$

As before, from the previous inequality we deduce that

$$d(\mu_N(E), \mu_N(\widetilde{E}_N^q)) \leqslant \frac{4}{k}\,.$$

Finally, we define the matrix B_N^q as

$$B_N^q = \begin{pmatrix} B_k^{(1)} & & & & & \\ & B_k^{(2)} & & & & \\ & & \ddots & & & \\ & & & B_k^{(q)} & & \\ & & & & 0_{r\times r}\,, \end{pmatrix}$$

where $B_k^{(i)} = \mathfrak{L}_k^{(i)} \widetilde{\mathcal{M}}_k^{(i)}$.

Let $\delta > 0$, for N large enough such that $\frac{4}{k} \leqslant \frac{\delta}{4}$, we have almost surely

$$d(\mu_N(C_N^q), \mu_N(\widetilde{C}_N^q)) + d(\mu_N(E), \mu_N(E_N^q)) > \frac{\delta}{2}.$$

As a consequence,

$$\begin{aligned}
&\mathbb{P}\left(d(\mu_N(C_N^q), \mu_N(E)) > \delta\right) \\
&\leqslant \mathbb{P}\left(d(\mu_N(C_N^q), \mu_N(\widetilde{C}_N^q)) + d(\mu_N(\widetilde{C}_N^q), \mu_N(B_N^q)) + d(\mu_N(B_N^q), \mu_N(E_N^q)) + d(\mu_N(E_N^q), \mu_N(E)) > \delta\right) \\
&\leqslant \mathbb{P}\left(d(\mu_N(\widetilde{C}_N^q), \mu_N(B_N^q)) + d(\mu_N(B_N^q), \mu_N(E_N^q)) > \frac{\delta}{2}\right)\,.
\end{aligned}$$

Moreover, combining Lemma 2.2 and Lemma 3.7 we deduce that

$$\begin{aligned}
d(\mu_N(\widetilde{C}_N^q), \mu_N(B_N^q)) &\leqslant \frac{2}{N}\sum_{i=1}^{q}\sum_{1\leqslant \ell,j\leqslant k} |\mathfrak{L}_k^{(i)}(\ell,j) - \mathcal{L}_k^{(i)}(\ell,j)|\,, \\
d(\mu_N(B_N^q), \mu_N(\widetilde{E}_N^q)) &\leqslant \frac{2}{N}\sum_{i=1}^{q}\sum_{1\leqslant \ell,j\leqslant k} |\mathfrak{M}_k^{(i)}(\ell,j) - \widetilde{\mathcal{M}}_k^{(i)}(\ell,j)|\,.
\end{aligned}$$

Applying Lemma 3.8 point $i.$, we deduce that

$$d(\mu_N(\widetilde{C}_N^q), \mu_N(B_N^q)) + d(\mu_N(B_N^q), \mu_N(\widetilde{E}_N^q)) \leqslant \frac{8}{N}\sum_{i=1}^{q}\sum_{j=0}^{k-1} Z^{(i)}_{\frac{k-j}{N}}\,,$$

where the last sum denotes the sum of independent random variables with law $Z_{\frac{k-j}{N}}$, defined in Lemma 3.8.

Thus, for N large enough such that $\frac{4}{k} \leqslant \frac{\delta}{4}$, we deduce that for any non-negative function $a(q^{-1})$:

$$\begin{aligned}
\mathbb{P}\left(d(\mu_N(C_N^q), \mu_N(E)) > \delta\right) &\leqslant \mathbb{P}\left(\sum_{i=1}^{q}\sum_{j=0}^{k-1} Z^{(i)}_{\frac{k-j}{N}} > \frac{N\delta}{16}\right) \\
&\leqslant e^{-a(q^{-1})N\delta/16}\left(\sup_{0<h<1}\mathbf{E}\left[\exp(a(h)Z_h)\right]\right)^{qk}\,,
\end{aligned}$$

Where in the last inequality we used Remark 3.9, namely, since $\frac{k-j}{N} \leqslant \frac{1}{q}$, we have

$$\mathbb{E}[\exp(a(q^{-1})Z_{\frac{k-j}{N}}] \leqslant \mathbb{E}[\exp(a(q^{-1})Z_{\frac{1}{q}}].$$

Setting $a(h) = -\frac{1}{2}\ln(h) + 1$ and applying Lemma 3.8 point *ii.*, we deduce that there exist constants $\widetilde{K}$ and $c > 0$, independent of q, such that

$$\frac{1}{N}\ln\left(\mathbb{P}\left(d(\mu_N(C_N^q), \mu_N(E)) > \delta\right)\right) \leqslant -c\ln(q)\delta + \widetilde{K}\,,$$

And we obtain the claim. □

We can apply the previous Lemma to study the case of continuous potential, indeed as a consequence of Varadhan's Lemma we obtain the main result of this section, namely

rnative LDP **Theorem 3.11.** *In the same notation as before. Let $\beta > 0$, and $V : \mathbb{T} \to \mathbb{R}$ continuous. The law of the empirical measures $\mu_N(E)$ under $\mathrm{d}\mathbb{P}_{C,N}^{V,\frac{2\beta}{N}}$ satisfies a large deviations principle at speed N, with a good rate function $I_\beta^V(\mu) = f_\beta^V(\mu) - \inf_{\nu\in\mathcal{P}(\mathbb{T})} f_\beta^V(\nu)$, where*

$$f_\beta^V(\mu) = \lim_{\delta\to 0}\liminf_{q\to\infty} \inf_{\substack{\nu_{\beta/q},\dots,\nu_\beta \\ \frac{1}{q}\sum_i \nu_{i\beta/q}\in B_\mu(\delta)}} \left\{\frac{1}{q}\sum_{i=1}^{q}\left(J_{i\beta/q}(\nu_{i\beta/q}) + \int_{\mathbb{T}} V\mathrm{d}\nu_{i\beta/M}\right)\right\}, \tag{3.10}$$

eq: rate fu

4 Proof of the main results

sec main thm

In this section, we conclude the proof of Theorem 1.4 and prove Theorem 1.5. The main tool to prove these theorems is the uniqueness of the minimizer of the rate function for the β ensemble.

Define the free energies of the Ablowtiz-Ladik lattice and the Circular beta ensemble at high temperature as

$$F_{AL}(V,\beta) = \inf_{\nu\in\mathcal{P}(\mathbb{T})} g_\beta^V(\nu)\,, \quad F_C(V,\beta) = \inf_{\nu\in\mathcal{P}(\mathbb{T})} f_\beta^V(\nu)\,,$$

where g_β^V, and f_β^V are given by (2.4) and (3.10). We claim that

n_conclusion **Lemma 4.1.** *Let $\beta > 0$, and $V : \mathbb{T} \to \mathbb{R}$ continuous, then the following holds:*

a. the map $\beta \to F_C(V,\beta)$ is continuously differentiable on $\mathbb{R}_+^*$. Moreover:

$$F_{AL}(V,\beta) = \partial_\beta(\beta F_C(V,\beta))\,;$$

b. for almost all $\beta > 0$ there exists a unique minimizer ν_β^V of the functional $J_\beta^V(\mu)$, see Corollary 2.5, given by

$$\nu_\beta^V = \partial_\beta(\beta\mu_\beta^V)\,,$$

i.e for continuous $f : \mathbb{T} \to \mathbb{R}$,

$$\int_{\mathbb{T}} f\mathrm{d}\nu_\beta^V = \partial_\beta\left(\beta\int_{\mathbb{T}} f\mathrm{d}\mu_\beta^V\right).$$

we recall that the measure μ_β^V is defined as the unique minimizer of the functional I_β^V in Theorem 3.11.

Remark 4.2. *Our definition of Free Energy is different from the one used in [24, 28]. Indeed, in virtue of Varadhan's Lemma [13, Theorem 1.2.1], we have*

$$\begin{aligned} F_{AL}(V,\beta) &= \inf_{\nu\in\mathcal{P}(\mathbb{T})} g_\beta^V(\nu) = -\lim_N \frac{1}{N}\ln \mathbb{E}^\beta_{AL,N}\left[e^{-TrV(\mathcal{E})}\right], \\ F_C(V,\beta) &= \inf_{\nu\in\mathcal{P}(\mathbb{T})} f_\beta^V(\nu) = -\lim_N \frac{1}{N}\ln \mathbb{E}^\beta_{C,N}\left[e^{-TrV(E)}\right], \end{aligned} \tag{4.1}$$

eq:equality

instead in [24, 28], the authors defined the free energies as

$$\begin{aligned} \widetilde{F}_{AL}(V,\beta) &= -\lim_{N\to\infty}\frac{1}{N}\ln(Z_N^{AL}(V,\beta)), \\ \widetilde{F}_C(V,\beta) &= -\lim_{N\to\infty}\frac{1}{N}\ln(Z_N^{C}(V,\beta)). \end{aligned}$$

We notice that it is possible to recover one expression from the other since

$$\begin{aligned} F_{AL}(V,\beta) &= \widetilde{F}_{AL}(V,\beta) - \widetilde{F}_{AL}(0,\beta), \\ F_C(V,\beta) &= \widetilde{F}_C(V,\beta) - \widetilde{F}_C(0,\beta). \end{aligned}$$

To prove uniqueness of the minimizer ν_β^V, we need to consider a continuous family $(\mu_s^*)_{0<s<\beta}$, where each μ_s^* minimizes J_s^V, see Corollary 2.5. We address the existence of such a family in the next Lemma, which we prove in the appendix A.

_minimizers **Lemma 4.3.** *Let $M_\beta^V = (J_\beta^V)^{-1}(\{0\})$ be the set of minimizers of J_β^V, defined in Corollary 2.5. Then, $\beta \mapsto M_\beta^V$ is continuous in the sense that for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $0 \leqslant h \leqslant \delta$, $M_{\beta+h}^V \subset (M_\beta^V)^\varepsilon$, where for $A \subset \mathcal{P}(\mathbb{T})$ we denote $A^\varepsilon = \{\mu \in \mathcal{P}(\mathbb{T}) \mid d(\mu, A) \leqslant \varepsilon\}$.*

Proof of Lemma 4.1. First, we notice that for any probability measure $\mu \in \mathcal{P}(\mathbb{T})$, Theorem 3.11 implies

$$\begin{aligned} f_\beta^V(\mu) &\geqslant \liminf_{q\to\infty} \inf_{\nu\in\mathcal{P}(\mathbb{T})} \left\{ \frac{1}{q}\sum_{i=1}^q \left(J_{i\beta/q}(\nu) + \int_{\mathbb{T}} V \mathrm{d}\nu \right) \right\} \\ &= \int_0^1 \inf_{\nu\in\mathcal{P}(\mathbb{T})} g_{s\beta}^V(\nu) = \int_0^1 F_{AL}(V, s\beta)\mathrm{d}s, \end{aligned} \tag{4.2}$$

eq:lb_energy

Where we noticed that the Riemann sums indeed converge towards the integral since $s \mapsto F_{\mathrm{AL}}(V, s\beta)$ is concave, this can be seen by applying Hölder inequality to equation (4.1).

To prove the first part of the claim, we show that the lower bound is achieved. For $s \in [0,1]$, let $\nu_{s\beta}^*$ be a minimizer of $\inf_{\nu\in\mathcal{P}(\mathbb{T})} g_{s\beta}^V(\nu)$. From Lemma 4.3, we can choose $\nu_{s\beta}^*$ such that the map $s \to \nu_{s\beta}^*$ is continuous. This implies that $\mu_\beta^* = \int_0^1 \nu_{s\beta}^* \mathrm{d}s$ is a well-defined probability measure on $\mathbb{T}$. We claim that this measure minimizes f_β^V (3.10), and so I_β^V. Indeed, from Theorem 3.11, we deduce that

$$\begin{aligned} f_\beta^V(\mu_\beta^*) &= \lim_{\delta\to 0}\liminf_{q\to\infty} \inf_{\substack{\nu_{\beta/q},\dots,\nu_\beta \\ \frac{1}{q}\sum_i \nu_{i\beta/q}\in B_{\mu_\beta^*}(\delta)}} \left\{ \frac{1}{q}\sum_{i=1}^q \left(J_{i\beta/q}(\nu_{i\beta/q}) + \int_{\mathbb{T}} V\mathrm{d}\nu_{i\beta/q} \right) \right\} \\ &\leqslant \liminf_{q\to\infty} \left\{ \frac{1}{q}\sum_{i=1}^q \left(J_{i\beta/q}(\nu_{i\beta/q}^*) + \int_{\mathbb{T}} V\mathrm{d}\nu_{i\beta/q}^* \right) \right\} \\ &= \liminf_{q\to\infty} \left\{ \frac{1}{q}\sum_{i=1}^q \inf_{\nu\in\mathcal{P}(\mathbb{T})} \left(J_{i\beta/q}(\nu) + \int_{\mathbb{T}} V\mathrm{d}\nu \right) \right\} \\ &= \int_0^1 \inf_{\nu\in\mathcal{P}(\mathbb{T})} g_{s\beta}^V(\nu) = \int_0^1 F_{AL}(V,s\beta)\mathrm{d}s. \end{aligned} \tag{4.3}$$

eq:ub_energy

Combining (4.2)-(4.3), and performing the change of coordinates $s\beta = t$ we deduce that:

$$\beta F_C(V,\beta) = \int_0^\beta F_{AL}(V,t)\mathrm{d}t\,.$$

Moreover, from Lemma 3.2 we deduce that the map $\beta \to F_C(V,\beta)$ is Lipschitz in β, and so almost surely differentiable. This implies that for almost all $\beta > 0$

$$F_{AL}(V,\beta) = \partial_\beta(\beta F_C(V,\beta))\,.$$

Furthermore, we have just shown that $I_\beta^V(\mu) = f_\beta^V(\mu) - \inf_{\nu\in\mathcal{P}(\mathbb{T})} f_\beta^V(\nu)$ reaches its minimum at $\int_0^1 \nu_{s\beta}^*\mathrm{d}s$. By uniqueness of the minimizer of $I_\beta^V(\mu)$, Theorem 3.1, we deduce that we have the equality between probability measures $\mu_\beta^V = \int_0^1 \nu_{s\beta}^*\mathrm{d}s$. Taking $f : \mathbb{T} \to \mathbb{R}$ continuous we get

$$\beta\int_{\mathbb{T}} f\mathrm{d}\mu_\beta^V = \int_0^\beta\int_{\mathbb{T}} f\mathrm{d}\nu_s^*\mathrm{d}s.$$

Note that the function $s \mapsto \int_{\mathbb{T}} f\mathrm{d}\nu_s^*$ is continuous, therefore by differentiating this equality, we get that ν_β^* is the unique minimizer of J_β^V, which we denote by ν_β^V, and satisfies for f continuous

$$\int_{\mathbb{T}} f\nu_\beta^V = \partial_\beta\left(\beta\int_{\mathbb{T}} f\mu_\beta^V\right)\,,$$

proving point b.

□

Remark 4.4. *As a corollary of the previous Lemma we obtain Theorem 1.4.*

5 The Schur Flow

sec: Schur

In this section, we consider another integrable model, namely the Schur flow. Our goal is to show that is possible to obtain a similar result to the one that we presented for the Ablowitz-ladik lattice. Namely, we prove the existence of a large deviations principle for the Schur flow, and we relate its density of state to the one of the Jacobi beta ensemble in the high temperature regime.

5.1 Generalized Gibbs Ensemble

The Schur flow is the system of ODEs [18]

$$\dot{\alpha_j} = \rho_j^2(\alpha_{j+1} - \alpha_{j-1})\,, \quad \rho_j = \sqrt{1-|\alpha_j|^2} \tag{5.1}$$

eq:schur

and, as before, we consider periodic boundary conditions, namely $\alpha_j = \alpha_{j+N}$ for all $j \in \mathbb{Z}$.

In [3], it is argued that the continuum limit of (5.1) is the modified Korteweg-de Vries equation:

$$\partial_t u = \partial_x^3 u - 6u^2\partial_x u\,.$$

We notice that, if one chooses an initial data such that $\alpha_j(0) \in \mathbb{R}$ for all $j = 1,\dots,N$, then $\alpha_j(t) \in \mathbb{R}$ for all times. Moreover, it is straightforeward to verify that $K_0 = \prod_{j=1}^N \left(1-|\alpha_j|^2\right)$ is conserved along the Schur flow. This implies that we can choose as phase space for the Schur flow the N-cube $\mathbb{I}^N$, where $\mathbb{I} := (-1,1)$.

On this phase space, we consider the Poisson braket (1.2), so we can rewrite the Schur flow (5.1) in Hamiltonian form as

$$\dot{\alpha}_j = \{\alpha_j, H_S\}, \quad H_S = -i \sum_{j=1}^{N} (\alpha_j \overline{\alpha}_{j+1} - \overline{\alpha}_j \alpha_{j+1}) \ .$$

It is well known that the Schur flow admits as Lax matrix the same one as the AL [18], namely $\mathcal{E}$ (1.4) is the Lax matrix of the Schur flow. This implies that the Ablowitz-Ladik's constants of motion are conserved also along the Schur flow (5.1).

Following the same construction made for the Ablowitz-Ladik lattice, on $\mathbb{I}^N$ we define the finite volume limit GGE as

$$\mathrm{d}\mathbb{P}^{V,\beta}_{S,N}(\alpha_1, \dots, \alpha_N) = \frac{1}{Z^S_N(V,\beta)} \prod_{j=1}^{N} (1-\alpha_j^2)^{\beta-1} \mathbf{1}_{\{\alpha_j \in \mathbb{I}\}} \exp(-\mathrm{Tr}(V(\mathcal{E}))) \mathrm{d}\boldsymbol{\alpha}, \tag{5.2}$$

GGE Schur

where $Z^S_N(V,\beta)$ is the partition function of the system

$$Z^S_N(V,\beta) = \int_{\mathbb{I}^N} \prod_{j=1}^{N} (1-\alpha_j^2)^{\beta-1} \exp(-\mathrm{Tr}(V(\mathcal{E}))) \mathrm{d}\boldsymbol{\alpha}.$$

Since according to the measure (5.2) the matrix $\mathcal{E}$ is real, its eigenvalues come in pairs [27], meaning that if $e^{i\theta_j}$ is an eigenvalue, then its conjugate $e^{-i\theta_j}$ is also an eigenvalue. This implies that for a system of size N even, there are just $n = N/2$ independent eigenvalues. Following the same idea as in [21], it is more convenient to restrict the argument of the eigenvalues in $[0, \pi)$ and then consider $x_j = \cos(\theta_j)$, $j = 1, \dots, n$. In these variables, the empirical spectral measure $\mu_n(\mathcal{E})$ reads:

$$\mu_n(\mathcal{E}) = \frac{1}{n} \sum_{j=1}^{n} \delta_{x_j} \ , \quad x_j \in \mathbb{I} \, . \tag{5.3}$$

eq:empirica

As a corollary of Lemma 2.3 and Proposition 2.4, we obtain the existence of a large deviations principle for the sequence $(\mu_n(\mathcal{E}))$, namely:

Corollary 5.1. *Let $V : \mathbb{I} \to \mathbb{R}$ be continuous. Under $\mathbb{P}^{V,\beta}_{S,n}$ the sequence $(\mu_n(\mathcal{E}))$ fulfils a large deviations principle with good, convex rate function $S^V_\beta(\mu) = h^V_\beta(\mu) - \inf_{\nu \in \mathcal{P}(\mathbb{I})} h^V_\beta(\nu)$, where*

$$h^V_\beta(\nu) = K_\beta(\nu) + \int_{\mathbb{I}} V \mathrm{d}\nu \, ,$$

where $K_\beta(\nu)$ is the rate function of μ_n under the law $\mathbb{P}^{0,\beta}_{S,n}$.

5.2 Jacobi beta ensemble in the high temperature regime

The Jacobi beta ensemble refers to the distribution of charges constrained to the segment $\mathbb{I}$, and subjected to an external potential $W(x) = -a \ln(1-x) - b \ln(1+x) + V(x)$, here $a, b > -1$ and $W(x) \in C^0(\mathbb{I})$. Specifically the joint distribution of these particles is

$$\mathrm{d}\mathbb{P}^{(V,\widetilde{\beta})}_{J,n} = \frac{1}{Z^J_N(V,\widetilde{\beta})} \prod_{i<j} |x_i - x_j|^{\widetilde{\beta}} \prod_{j=1}^{n} (1-x_j)^a (1+x_j)^b e^{-V(x_j)} \mathrm{d}x_j \, . \tag{5.4}$$

eq:JbE

In [21], Killip and Nenciu were able to show that the distribution (5.4) can be realized as the eigenvalues distribution of a particular CMV matrix, specifically they proved the following

Theorem 5.2 (cf. [21] Proposition 5.3)**.** *Let $N = 2n$, consider the CMV matrix E in* (3.4) *with parameters $\alpha_1, \dots, \alpha_{2n-1} \in \mathbb{I}$ distributed according to*

$$\mathrm{d}\mathbb{B}_n^{(V,\widetilde{\beta})} = \frac{1}{\mathfrak{Z}_n(V,\widetilde{\beta})} \prod_{j=1}^{2n-1} (1-\alpha_j^2)^{\widetilde{\beta}(2n-j)/4-1} \prod_{j=1}^{2n-1} (1-\alpha_j)^{a+1-\widetilde{\beta}/4} (1+(-1)^{j+1}\alpha_j)^{b+1-\widetilde{\beta}/4} e^{TrV(E)} \mathrm{d}\alpha_j \,,$$

and $\alpha_{2n} = -1$, here $\mathfrak{Z}_N(V,\beta)$ is the normalization constant. Then all the eigenvalues of E come in pairs, meaning that if $e^{i\theta_j}$ is an eigenvalue, then also $e^{-i\theta_j}$ is one. Moreover, under the change of variables $\cos(\theta_j) = x_j$, the x_js are distributed according to (5.4).

Remark 5.3. *We notice that the previous proposition is not stated in this way in [21], but this equivalent formulation is more useful for our purpose.*

Also in this case, we are interested in the high temperature regime for this ensemble. Specifically we consider the situation $\widetilde{\beta} = \frac{4\beta}{N} = \frac{2\beta}{n}$, and $a = b = -1 + \frac{\widetilde{\beta}}{4}$, in this regime $\mathrm{d}\mathbb{P}_{J,n}^{(V,\frac{\beta}{n})}$ reads

$$\mathrm{d}\mathbb{P}_{J,n}^{(V,\frac{2\beta}{n})} = \frac{1}{Z_N^J\left(V,\frac{\beta}{n}\right)} \prod_{i<j} |x_i - x_j|^{\frac{2\beta}{n}} \prod_{j=1}^{n} (1-x_j)^{-1+\frac{\beta}{2n}} (1+x_j)^{-1+\frac{\beta}{2n}} e^{-V(x_j)} \mathrm{d}x_j \,, \tag{5.5}$$

eq:JbE_ht

and $\mathrm{d}\mathbb{B}_n^{(V,\frac{\beta}{n})}$ becomes

$$\mathrm{d}\mathbb{B}_n^{(V,\frac{\beta}{n})} = \frac{1}{\mathfrak{Z}_n\left(V,\frac{\beta}{n}\right)} \prod_{j=1}^{2n-1} (1-\alpha_j^2)^{\beta\left(1-\frac{j}{2n}\right)-1} \prod_{j=1}^{2n-1} e^{\mathrm{Tr}V(E)} \mathrm{d}\alpha_j \,.$$

We mention that this particular regime was considered in [15,33]. In these papers the authors computed the density of states for this ensemble in the case $V = 0$.

We can apply [16, Corollary 1.3] to (5.5) to obtain a large deviations principle for the empirical measure $\mu_n(E) = \frac{1}{n}\sum_{j=1}^n \delta_{x_j}$. Specifically, we deduce that

Proposition 5.4. *For any continuous $V : \mathbb{I} \to \mathbb{R}$. The law of the empirical measures $\mu_n(E)$ under $\mathrm{d}\mathbb{P}_{J,n}^{(V,\frac{2\beta}{n})}$ satisfies a large deviations principle at speed n in the space $\mathcal{P}(\mathbb{I})$, with a good rate function $\mu \mapsto Q_\beta^V(\mu)$ given for μ absolutely continuous with respect to Lebesgue measure, and with density $\frac{\mathrm{d}\mu}{\mathrm{d}x}$, by $Q_\beta^V(\mu) = q_\beta^V(\mu) - \inf_{\nu\in\mathcal{P}(\mathbb{I})} q_\beta^V(\nu)$, where*

$$q_\beta^V(\mu) = \int_{\mathbb{I}} (V(x)+\ln(1+x)+\ln(1-x))\mathrm{d}\mu(x) - 2\beta \int_{\mathbb{I}\times\mathbb{I}} \ln(|x-y|)\mathrm{d}\mu(x)\mathrm{d}\mu(y) + \int_{\mathbb{I}} \ln\left(\frac{\mathrm{d}\mu}{\mathrm{d}x}(x)\right) \mathrm{d}\mu(x) \,, \tag{5.6}$$

eq: functi

and $Q_\beta^V(\mu) = +\infty$ otherwise.

We notice that the arguments in Section 3 and 4 can be applied also in this context with $\mathrm{d}\mathbb{P}_{J,n}^{(V,\frac{2\beta}{n})}$ in place of $\mathrm{d}\mathbb{P}_{C,N}^{(V,\frac{2\beta}{N})}$, and $\mathrm{d}\mathbb{P}_{S,N}^{V,\beta}$ in place of $\mathrm{d}\mathbb{P}_{AL,N}^{V,\beta}$. Hence, we deduce the following result

Theorem 5.5. *Consider the sequence of measures $\mu_n(\mathcal{E})$* (5.3) *under the law $\mathrm{d}\mathbb{P}_{S,2n}^{V,\beta}$* (5.2)*, then*

$$\mu_n(\mathcal{E}) \xrightarrow{a.s.} \nu_\beta^V \,.$$

Moreover, ν_β^V is absolutely continuous with respect to the Lebesgue measure, and it reads

$$\nu_\beta^V = \partial_\beta(\beta\mu_\beta^V) \,,$$

where μ_β^V is the unique minimizer of the functional q_β^V (5.6)*.*

Finally, it is worth to mention that in the case $V(x) = 0$, it is possible to compute explicitly the densities of states for both the Jacobi beta ensemble at high temperature and for the Schur flow [15, 23].

A Technical Results

app:A

In this appendix we collect the proof of all the technical results that we exploit along the proof of the main theorem. For reader convenience, we report here the statement of Lemmas.

Proof of Lemma 2.2

Lemma A.1. *For any A, B unitary matrices of size $N \times N$, we have*

- *For f with bounded variation,*

$$\left| \int f d\mu - \int f d\nu \right| \leqslant \|f\|_{BV} \frac{rank(A-B)}{N},$$

- *For f Lipschitz,*

$$\left| \int f d\mu - \int f d\nu \right| \leqslant \|f\|_{Lip} \frac{1}{N} \sum_{i,j} |(A-B)_{i,j}|.$$

As a consequence,

$$d(\mu(A), \mu(B)) \leqslant \min \left\{ \frac{rank(A-B)}{N}, \frac{1}{N} \sum_{i,j} |(A-B)_{i,j}| \right\}. \qquad \text{distance}$$

Proof. The first point is a consequence of the fact that the eigenvalues of A and B interlace on the unit circle.

First, we order the eigenvalues $\lambda_1(A), \ldots, \lambda_N(A), \lambda_1(B), \ldots, \lambda_N(B)$ of A, B in such a way that

$$-\pi \leqslant \arg(\lambda_1(A)) \leqslant \ldots \leqslant \arg(\lambda_N(A)) < \pi\,,$$

and analogously for B.

Write $B = (I_N + (B-A)A^{-1})A$ and set $U := I_N + (B-A)A^{-1}$. One checks that U is unitary, $B = UA$, and that $\mathrm{rank}(U - I) = \mathrm{rank}(B - A) =: r$. By [5, section 6, equation (85)], we deduce that for $1 \leqslant j \leqslant N$

$$\arg(\lambda_{j-r}(A)) \leqslant \arg(\lambda_j(B)) \leqslant \arg(\lambda_{j+r}(A))\,. \qquad \text{(A.1)} \qquad \text{interlacing}$$

This means that $\lambda_j(B)$ lies on the anticlockwise arc $(\arg(\lambda_{j-r}(A)), \arg(\lambda_{j+r}(A)))$ of the circle. If $j - r \leqslant 0$ we identify λ_{j-r} with λ_{j-r+N}, and analogously for the case $j + r > N$.
It is a classical result (see [4]) to deduce from (A.1) that

$$\left| \int f d\mu_N(A) - \int f d\mu_N(B) \right| \leqslant \|f\|_{\mathrm{BV}} \frac{r}{N} = \frac{r}{N}\,,$$

for any $f : \mathbb{T} \to \mathbb{R}$ such that $||f||_{BV} \leqslant 1$. As a consequence, we obtain the first point.
The proof of the second point is the same as in the symmetric case, see [19, (16)]. Indeed, we only use the fact that a normal matrix is unitarily diagonalizable. □

Proof of Lemma 3.7

Lemma A.2. *Let $N = 2k$ be even and A be a $N \times N$ matrix. Then,*

- $\sum_{i,j} |(\mathcal{L}A)_{i,j}| \leqslant 2 \sum_{i,j} |A_{i,j}|$,
- $\sum_{i,j} |(A\mathcal{M})_{i,j}| \leqslant 2 \sum_{i,j} |A_{i,j}|$,

where $\mathcal{M}$, and $\mathcal{L}$ are defined in (1.3).

Proof. We will just prove the first point, since the proof of the second one follows the same lines.
For $0 \leqslant l \leqslant k-1$ and $1 \leqslant j \leqslant N$, consider

$$(\mathcal{L}A)_{2l+1,j} = \overline{\alpha}_{2l+1} A_{2l+1,j} + \rho_{2l+1} A_{2l+2,j} \quad \text{and} \quad (\mathcal{L}A)_{2l+2,j} = \rho_{2l+1} A_{2l+1,j} - \alpha_{2l+2} A_{2l+2,j}.$$

Summing over i, j,

$$\sum_{i,j} |(\mathcal{L}A)_{i,j}| = \sum_{l=0}^{k-1} \sum_{j=1}^{N} |(\mathcal{L}A)_{2l+1,j}| + |(\mathcal{L}A)_{2l+2,j}| \leqslant 2 \sum_{l=0}^{k-1} \sum_{j=1}^{N} |A_{2l+1,j}| + |A_{2l+2,j}| = 2 \sum_{i,j} |A_{i,j}|,$$

where we used that $|\alpha_i|, \rho_i \leqslant 1$. □

Proof of Lemma 3.8

Lemma A.3. *Let α_ν and $\alpha_{\nu+h}$ defined by equation* (3.6). *Define $\rho_\nu = \sqrt{1 - |\alpha_\nu|^2}$, and $\rho_{\nu+h} = \sqrt{1 - |\alpha_{\nu+h}|^2}$, then the following hold*

i.

$$|\alpha_\nu - \alpha_{\nu+h}| \leqslant \frac{Y_h}{(X_1^2 + X_2^2 + Y_h^2)^{\frac{1}{2}}}, \text{ almost surely},$$
$$|\rho_\nu - \rho_{\nu+h}| \leqslant \frac{Y_h}{(X_1^2 + X_2^2 + Y_h^2)^{\frac{1}{2}}}, \text{ almost surely},$$

where $X_1, X_2 \sim \mathcal{N}(0,1)$, $Y_h \sim \chi_h$ are all independent.

ii. define $Z_h = \frac{Y_h}{(X_1^2+X_2^2+Y_h^2)^{\frac{1}{2}}}$, and $a(h) = -\frac{1}{2}\ln(h) + 1$, then there exists a constant K independent of h such that

$$\sup_{0<h<1} \mathbf{E}\left[\exp(a(h) Z_h)\right] \leqslant K. \tag{A.2}$$

eq:sup_boun

Proof. First, we focus on claim *i.*. We recall that $\alpha_\nu, \alpha_{\nu+h}$ are defined by

$$\alpha_\nu = \frac{X_1 + iX_2}{(X_1^2 + X_2^2 + Y_{\nu-1}^2)}, \quad \alpha_{\nu+h} = \frac{X_1 + iX_2}{(X_1^2 + X_2^2 + Y_{\nu-1}^2 + Y_h^2)}.$$

From the previous equation, we deduce that

$$\begin{aligned}
|\alpha_\nu - \alpha_{\nu+h}| &= \frac{|X_1 + iX_2|}{(X_1^2 + X_2^2 + Y_{\nu-1}^2)^{\frac{1}{2}}} \left(1 - \left(\frac{X_1^2 + X_2^2 + Y_{\nu-1}^2}{X_1^2 + X_2^2 + Y_{\nu-1}^2 + Y_h^2}\right)^{\frac{1}{2}}\right) \\
&= \frac{|X_1 + iX_2|}{(X_1^2 + X_2^2 + Y_{\nu-1}^2)^{\frac{1}{2}}} \left(1 - \left(1 - \frac{Y_h^2}{X_1^2 + X_2^2 + Y_{\nu-1}^2 + Y_h^2}\right)^{\frac{1}{2}}\right) \\
&\leqslant \left(\frac{X_1^2 + X_2^2}{X_1^2 + X_2^2 + Y_{\nu-1}^2}\right)^{\frac{1}{2}} \frac{Y_h}{(X_1^2 + X_2^2 + Y_{\nu-1}^2 + Y_h^2)^{\frac{1}{2}}},
\end{aligned}$$

where we used in the previous line that for $0 \leqslant a \leqslant b$ we have

$$\sqrt{b} \leqslant \sqrt{b-a} + \sqrt{a} \tag{A.3}$$ eq_sqroot

and we took $a = \frac{Y_h^2}{X_1^2+X_2^2+Y_{\nu-1}^2+Y_h^2}$, $b = 1$. The last term is bounded by the announced bound.

One can proceed analogously for $|\rho_\nu - \rho_{\nu+h}|$ obtaining that

$$\begin{aligned}|\rho(\alpha_{\nu+h}) - \rho(\alpha_\nu)| &= \sqrt{1-|\alpha_{\nu+h}|^2} - \sqrt{1-|\alpha_\nu|^2} \leqslant \sqrt{|\alpha_\nu|^2 - |\alpha_{\nu+h}|^2} \\ &= \sqrt{\frac{X_1^2+X_2^2}{X_1^2+X_2^2+Y_{\nu-1}^2}}\sqrt{1-\frac{X_1^2+X_2^2+Y_{\nu-1}^2}{X_1^2+X_2^2+Y_{\nu-1}^2+Y_h^2}} \leqslant \frac{Y_h}{(X_1^2+X_2^2+Y_{\nu-1}^2+Y_h^2)^{\frac{1}{2}}}\end{aligned}$$

where we used again equation (A.3) with $a = 1-|\alpha_\nu|^2$ and $b = 1-|\alpha_{\nu+h}|^2$. Thus, point *i.* is proved.

To prove point *ii.*, we find explicitly the law of Z_h. Thus, we consider a continuous function $f:\ (0,1) \to \mathbb{R}$, and we compute:

$$\int_{\mathbb{R}^2\times\mathbb{R}_+} f\left(\frac{y}{\left(x_1^2+x_2^2+y^2\right)^{\frac{1}{2}}}\right) e^{-\frac{x_1^2+x_2^2+y^2}{2}} y^{h-1}\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y\,.$$

Performing the change of coordinates $(u,v) = \frac{1}{(x_1^2+x_2^2+y^2)^{1/2}}(x_1,x_2)$, which is the same one that we performed in Lemma 3.5, we obtain that

$$\begin{aligned}&\int_{\mathbb{R}^2\times\mathbb{R}_+} f\left(\frac{y}{\left(x_1^2+x_2^2+y^2\right)^{\frac{1}{2}}}\right) e^{-\frac{x_1^2+x_2^2+y^2}{2}} y^{h-1}\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y \\ &= \int_{\mathbb{D}\times\mathbb{R}_+} \frac{f\left(\sqrt{1-u^2-v^2}\right)}{(1-u^2-v^2)^2} e^{-\frac{y^2}{2(1-u^2-v^2)}} y^{h+1}\mathrm{d}u\mathrm{d}v\mathrm{d}y \\ &\overset{\sqrt{1-u^2-v^2}t=y}{=} \int_{\mathbb{D}\times\mathbb{R}_+} f\left(\sqrt{1-u^2-v^2}\right)\left(1-u^2-v^2\right)^{\frac{h}{2}-1} e^{-\frac{t^2}{2}} t^{h+1}\mathrm{d}u\mathrm{d}v\mathrm{d}t\,.\end{aligned}$$

We can now explicitly compute the integral in t. Moreover, we can express the remaining part of the integral in polar coordinates; namely, we apply the change of variables $u = \rho\cos(\theta), v = \rho\sin(\theta)$, obtaining that:

$$\begin{aligned}&\int_{\mathbb{D}\times\mathbb{R}_+} f\left(\sqrt{1-u^2-v^2}\right)\left(1-u^2-v^2\right)^{\frac{h}{2}-1} e^{-\frac{t^2}{2}} t^{h+1}\mathrm{d}u\mathrm{d}v\mathrm{d}t \\ &= 2\pi 2^{\frac{h}{2}}\Gamma\left(\frac{h}{2}+1\right)\int_0^1 \rho f(\sqrt{1-\rho^2})\left(1-\rho^2\right)^{\frac{h}{2}-1}\mathrm{d}\rho \\ &\overset{\sqrt{1-\rho^2}=w}{=} 2\pi 2^{\frac{h}{2}}\Gamma\left(\frac{h}{2}+1\right)\int_0^1 f(w)w^{h-1}\mathrm{d}w\,,\end{aligned}$$

here $\Gamma(x)$ is the gamma function (3.5). Thus, in order to obtain the estimate (A.2), we have to deduce an upper bound for

$$\sup_{0<h<1} \frac{\int_0^1 e^{a(h)w}w^{h-1}\mathrm{d}w}{\int_0^1 w^{h-1}\mathrm{d}w}\,.$$

For any $0 < h < 1$, we can explicitly compute the denominator as

$$\int_0^1 w^{h-1}\mathrm{d}w = \frac{1}{h}\,. \tag{A.4}$$ eq:denomina

Moreover, we can give an upper bound on the numerator as

$$\begin{aligned}\int_0^1 e^{a(h)w} w^{h-1}\mathrm{d}w &\overset{a(h)w=r}{=} \frac{1}{a(h)^h}\int_0^{a(h)} e^r r^{h-1}\mathrm{d}r = \frac{1}{a(h)^h}\left(\int_0^1 e^r r^{h-1}\mathrm{d}r + \int_1^{a(h)} e^r r^{h-1}\mathrm{d}r\right)\\ &\leqslant \frac{1}{a(h)^h}\left(e\int_0^1 r^{h-1}\mathrm{d}r + \int_1^{a(h)} e^r\mathrm{d}r\right) \leqslant \frac{e}{a(h)^h h} + \frac{e^{a(h)}}{a(h)^h}\,.\end{aligned} \tag{A.5}$$ eq:numerato

Combining (A.4)-(A.5), with our choice of $a(h) = -\frac{1}{2}\ln(h) + 1$, we deduce that there exists a constant K independent of h such that (A.2) holds.

□

Proof of Lemma 4.3

Lemma A.4. *Let $M_\beta^V = (J_\beta^V)^{-1}(\{0\})$ be the set of minimizers of J_β^V. Then, $\beta \mapsto M_\beta^V$ is continuous in the sense that for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $0 \leqslant h \leqslant \delta$, $M_{\beta+h}^V \subset (M_\beta^V)^\varepsilon$, where for $A \subset \mathcal{P}(\mathbb{T})$ we denote $A^\varepsilon = \{\mu \in \mathcal{P}(\mathbb{T}) \mid d(\mu, A) \leqslant \varepsilon\}$.*

Proof. Let $\varepsilon > 0$. We are going to show that for $h > 0$ small enough, we have

$$-\inf_{\left[(M_\beta^V)^\varepsilon\right]^c} J_{\beta+h}^V < 0,$$

which will ensure that $J_{\beta+h}^V > 0$ on $\left[(M_\beta^V)^\varepsilon\right]^c$, thus $\left[(M_\beta^V)^\varepsilon\right]^c \subset \left[(M_{\beta+h}^V)\right]^c$, and hence the conclusion.
By the large deviations principle for $(\mu_N)_{N \text{ even}}$ under $\mathbb{P}_{AL,N}^{V,\beta}$, Corollary 2.5, since $\left[(M_\beta^V)^\varepsilon\right]^c$ is open, we have

$$\begin{aligned}-\inf_{\left[(M_\beta^V)^\varepsilon\right]^c} J_{\beta+h}^V &\leqslant \liminf_{N \text{ even}} \frac{1}{N}\ln \mathbb{P}_{AL,N}^{V,\beta+h}\left(\mu_N(\mathcal{E}) \in \left[(M_\beta^V)^\varepsilon\right]^c\right)\\ &= \liminf_{N \text{ even}} \frac{1}{N}\ln \mathbb{P}_{AL,N}^{V,\beta+h}\left(d(\mu_N(\mathcal{E}), M_\beta^V) > \varepsilon\right)\\ &\leqslant \limsup_{N \text{ even}} \frac{1}{N}\ln \mathbb{P}_{AL,N}^{V,\beta+h}\left(d(\mu_N(\mathcal{E}), M_\beta^V) \geqslant \varepsilon\right)\,.\end{aligned}$$

Since for any positive h and $\boldsymbol{\alpha} \in \mathbb{D}^N$ $\prod_{j=1}^N (1-|\alpha_j|^2)^h \leqslant 1$,we deduce that for any $A \subseteq \mathbb{D}^N$

$$\frac{1}{N}\ln\left(\mathbb{P}_{AL,N}^{V,\beta+h}(A)\right) \leqslant \frac{1}{N}\left(\ln\left(\frac{Z_N^{AL}(V,\beta)}{Z_N^{AL}(V,\beta+h)}\right) + \ln\left(\mathbb{P}_{AL,N}^{V,\beta}(A)\right)\right)\,,$$

we recall that $\mathbb{P}_{AL,N}^{V,\beta}$ is defined in (1.6).

Applying the previous inequality in the case $A = \{d(\mu_N(\mathcal{E}), M_\beta^V) \geqslant \varepsilon\}$, we conclude that

$$-\inf_{\left[(M_\beta^V)^\varepsilon\right]^c} J_{\beta+h}^V \leqslant \limsup_{N\to\infty} \frac{1}{N}\left(\ln\left(\frac{Z_N^{AL}(V,\beta)}{Z_N^{AL}(V,\beta+h)}\right) + \ln\left(\mathbb{P}_{AL,N}^{V,\beta}(d(\mu_N(\mathcal{E}), M_\beta^V) \geqslant \varepsilon)\right)\right)\,.$$

From Corollary 2.5, we deduce that there exists a positive constant c, independent of h, such that

$$\limsup_{N\to\infty}\frac{1}{N}\mathbb{P}^{V,\beta}_{AL,N}(d(\mu_N(\mathcal{E}),M^V_\beta)\geqslant\varepsilon))\leqslant-\inf_{\left[(M^V_\beta)^\varepsilon\right]^c}J^V_\beta<-c\,.$$

Thus, to conclude we have just to prove that the function $g(\beta)=\lim_{N\to\infty}\frac{1}{N}\ln\left(Z^{AL}_N(V,\beta)\right)$ is continuous in β. Actually, we prove that this function is convex in β. Let $1/p+1/q=1$, and $\beta_1,\beta_2\in\mathbb{R}_+$ then

$$\begin{aligned}Z^{AL}_N\left(V,\frac{\beta_1}{p}+\frac{\beta_2}{q}\right)&=\int_{\mathbb{D}^N}\prod_{j=1}^N(1-|\alpha_j|^2)^{\frac{\beta_1}{p}+\frac{\beta_2}{q}-1}\exp(-\mathrm{Tr}(V(\mathcal{E})))\mathrm{d}^2\boldsymbol{\alpha}\\&=\int_{\mathbb{D}^N}\prod_{j=1}^N(1-|\alpha_j|^2)^{\frac{\beta_1-1}{p}+\frac{\beta_2-1}{q}}\exp\left(-\left(\frac{1}{p}+\frac{1}{q}\right)\mathrm{Tr}(V(\mathcal{E}))\right)\mathrm{d}^2\boldsymbol{\alpha}\\&\leqslant Z^{AL}_N\left(V,\beta_1\right)^{\frac{1}{p}}Z^{AL}_N\left(V,\beta_2\right)^{\frac{1}{q}}\,,\end{aligned}$$

where in the last inequality we used Hölder inequality. This implies that

$$g\left(\frac{\beta_1}{p}+\frac{\beta_2}{q}\right)\leqslant\frac{1}{p}g(\beta_1)+\frac{1}{q}g(\beta_2)\,,$$

thus $g(\beta)$ is convex, and so continuous, for $\beta>0$. We can now choose h is such a way that

$$\left|\limsup_{N\to\infty}\frac{1}{N}\ln\left(\frac{Z^{AL}_N(V,\beta)}{Z^{AL}_N(V,\beta+h)}\right)\right|<c\,,$$

so we obtain that

$$\inf_{\left[(M^V_\beta)^\varepsilon\right]^c}J^V_{\beta+h}>0\,.$$

□

Acknoledgments
This material is based upon work supported by the National Science Foundation under Grant No. DMS-1928930 while the author participated in a program hosted by the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 20-21 semester "Universality and Integrability in Random Matrix Theory and Interacting Particle Systems". G.M. has received funding from the European Union's H2020 research and innovation programme under the Marie Skłodowska–Curie grant No. 778010 *IPaDEGAN*.

References

Ablowitz1973 [1] M. J. Ablowitz, D. J. Kaup, and A. C. Newell, *Coherent pulse propagation, a dispersive, irreversible phenomenon*, J. Math. Phys., 15 (1973), pp. 1852–1858.

Ablowitz1974 [2] M. J. Ablowitz and J. F. Ladik, *Nonlinear differential-difference equations*, J. Math. Phys., 16 (1974), pp. 598–603.

Ablowitz1975 [3] ——, *Nonlinear differential-difference equations and Fourier analysis*, J. Math. Phys., 17 (1975), pp. 1011–1018.

Alicebook [4] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, vol. 118 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2010.

ArbenzGolub [5] P. Arbenz and G. H. Golub, *On the spectral decomposition of hermitian matrices modified by low rank perturbations with applications*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 40–58.

Cafasso [6] M. Cafasso, *Matrix biorthogonal polynomials on the unit circle and non-abelian Ablowitz-Ladik hierarchy*, J. Phys. A, 42 (2009), pp. 365211, 20.

Cantero2005 [7] M. J. Cantero, L. Moral, and L. Velázquez, *Minimal representations of unitary operators and orthogonal polynomials on the unit circle*, Linear Algebra Appl., 408 (2005), pp. 40–65.

Deift [8] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, vol. 3 of Courant Lecture Notes in Mathematics, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

DZ [9] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, vol. 38 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

DLMF [10] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.1.3 of 2021-09-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

Doyon_notes [11] B. Doyon, *Lecture Notes On Generalised Hydrodynamics*, SciPost Phys. Lect. Notes, (2020), p. 18.

dued1 [12] I. Dumitriu and A. Edelman, *Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models*, J. Math. Phys., 47 (2006), pp. 063302, 36.

DupuisEllis [13] P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations*, Chichester: John Wiley & Sons, 1997.

Ercolani [14] N. M. Ercolani and G. Lozano, *A bi-Hamiltonian structure for the integrable, discrete non-linear Schrödinger system*, Physica D, 218 (2006), pp. 105–121.

rrester2021 [15] P. J. Forrester and G. Mazzuca, *The classical β-ensembles with β proportional to $1/N$: from loop equations to Dyson's disordered chain*, J. Math. Phys., 62 (2021), pp. Paper No. 073505, 22.

Garcia [16] D. García-Zelada, *A large deviation principle for empirical measures on Polish spaces: application to singular Gibbs measures on manifolds*, Ann. Inst. Henri Poincaré Probab. Stat., 55 (2019), pp. 1377–1401.

htmanNenciu [17] M. Gekhtman and I. Nenciu, *Multi-hamiltonian structurefor the finite defocusing ablowitz-ladik equation*, Comm. Pure Appl. Math., 62 (2009), pp. 147–182.

Golinskii [18] L. Golinskii, *Schur flows and orthogonal polynomials on the unit circle*, Sb. Math., 197 (2006), pp. 1145 – 1165.

GMToda [19] A. Guionnet and R. Memin, *Large deviations for gibbs ensembles of the classical toda chain*. Preprint available on arXiv:2103.04858, 2021.

Hardy2021 [20] A. Hardy and G. Lambert, *CLT for circular beta-ensembles at high temperature*, J. Funct. Anal., 280 (2021), pp. Paper No. 108869, 40.

Killip2004 [21] R. Killip and I. Nenciu, *Matrix models for circular ensembles*, Int. Math. Res. Not., (2004), pp. 2665–2701.

Kilnen07 [22] R. Killip and I. Nenciu, *CMV: the unitary analogue of Jacobi matrices*, Comm. Pure Appl. Math., 60 (2007), pp. 1148–1188.

uca2021mean [23] G. Mazzuca, *On the mean density of states of some matrices related to the beta ensembles and an application to the toda lattice.* preprint available on arXiv:2008.04604, 2021.

generalized [24] G. Mazzuca and T. Grava, *Generalized gibbs ensemble of the ablowitz-ladik lattice, circular β-ensemble and double confluent heun equation.* Preprint available on arXiv: 2107.02303.

Nenciu2005 [25] I. Nenciu, *Lax pairs for the Ablowitz-Ladik system via orthogonal polynomials on the unit circle*, Int. Math. Res. Not., 2005 (2005), pp. 647–686.

Nijhoff1 [26] G. Quispel, F. Nijhoff, H. Capel, and J. Van Der Linden, *Linear integral equations and nonlinear difference-difference equations*, Physica A: Statistical Mechanics and its Applications, 125 (1984), pp. 344–380.

Simon2005 [27] B. Simon, *Orthogonal Polynomials on the Unit Circle*, vol. 54.1 of Colloquium Publications, American Mathematical Society, Providence, Rhode Island, 2005.

ydrodynamic [28] H. Spohn, *Hydrodynamic equations for the ablowitz-ladik discretization of the nonlinear schroedinger equation.* Preprint available on arXiv: 2107.04866.

Spohn2 [29] ——, *Ballistic space-time correlators of the classical toda lattice*, J. Phys., A53 (2020), p. 265004.

Spohn3 [30] ——, *The collision rate ansatz for the classical toda lattice*, Phys Rev., E101 (2020), p. 060103.

Spohn1 [31] ——, *Generalized Gibbs Ensembles of the Classical Toda Chain*, J. Stat. Phys., 180 (2020), pp. 4–22.

Spohn4 [32] ——, *Hydrodynamic equations for the toda lattice*, ArXiv 2101.06528, (2021).

Trinh2021 [33] H. D. Trinh and K. D. Trinh, *Beta Jacobi Ensembles and Associated Jacobi Polynomials*, J. Stat. Phys., 185 (2021), p. Paper No. 4.

sianMoments [34] A. Winkelbauer, *Moments and absolute moments of the normal distribution*, arXiv:1209.4340v2, (2014).

ZS [35] V. Zkharov and A. Shabat, *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media*, Soviet Physics JETP, 34 (1972), pp. 62–69.