1. Introduction

In this paper we will develop a probabilistic/statistical mechanical approach for producing solutions to real Monge-Ampère equations in $\mathbb{R}^n$, satisfying the second boundary value problem with respect to a given target convex body $P$. This fits naturally into the theory of optimal transport [38, 39]. In particular, it will lead to a probabilistic construction of optimal transport plans from a set $X$ in $\mathbb{R}^n$ to a target convex body $P$. The approach arose as a “spin-off effect” of the authors work on a probabilistic approach to Kähler-Einstein metrics on complex algebraic varieties and, more generally, complex Monge-Ampère equations on complex manifolds (see [10] for an out-line of the general complex geometric setting and [3] for connections to emergent gravity and boson-fermion correspondences in physics). For example, the permamental random point processes on $\mathbb{R}^n$ introduced below, which are determined by the target convex body $P$, are the push-forwards to $\mathbb{R}^n$ of determinantal point-processes defined on the complex torus $\mathbb{C}^n$ and the corresponding limiting real Monge-Ampère measures on $\mathbb{R}^n$ are the push-forwards of the corresponding limiting complex Monge-Ampère measures on $\mathbb{C}^n$. The push-forward map is the one induced from the standard identification of $\mathbb{C}^n$ with $T^n \times \mathbb{R}^n$, where $T^n$ is the real unit-torus and in the case when the convex body $P$ is a rational polytope the corresponding determinantal point processes are naturally defined on the corresponding toric projective algebraic variety $X_P$, compactifying $\mathbb{C}^n$. From this point
of view the optimal transport theory in $\mathbb{R}^n$ thus arises as the “push-forward” of the pluripotential theory appearing in the complex setting [4 5 7].

The general complex geometric framework will be considered in detail elsewhere [11]. Accordingly, we will in this paper concentrate on the corresponding real setting (see however sections [6.4] and [6.9] for some relations to the complex setting).

1.1. The Monge-Ampère, optimal transport and permanental point processes.

In their simplest classical form the real Monge-Ampère equations that are the focus of the present paper are of the form

$$\det(\frac{\partial^2 \phi}{\partial x_i \partial x_j}) = e^{\beta \phi} \rho_0(x)$$

for a given function $\rho_0(x)$ of unit-mass and a given parameter $\beta \geq 0$. As usual the solution $\phi$ is demanded to be convex, but to ensure uniqueness further growth conditions at infinity have to be specified. The relevant situation here will be when $\phi$ grows as the support function $\phi_P$ of a given $n$-dimensional convex body $P$ in $\mathbb{R}^n$. In the PDE literature [1, 39] this is sometimes called the second boundary value problem for the equation above and (for $\rho_0$ strictly positive) it turns out to be equivalent to demanding that the gradient $\nabla \phi$ map $\mathbb{R}^n$ diffeomorphically onto the interior of $P$:

$$\nabla \phi : \mathbb{R}^n \to P$$

Accordingly, $P$ is sometimes referred to as the target. More generally, we will consider the setting of a given triple $(P, \mu_0, \phi_0)$ consisting of a convex body $P$ in $\mathbb{R}^n$ a (Borel) measure $\mu_0$, whose support will be denoted by $X$ and a weight function $\phi_0$ on $\mathbb{R}^n$, i.e. a (a possible non-convex) continuous function $\phi_0$ which grows faster than the support function $\phi_P$ of $P$ at infinity in $\mathbb{R}^n$ (see section 5.1.1) and such that $e^{-\beta \phi_0} \mu_0$ has finite total mass. The corresponding Monge-Ampère equation is then

$$MA(\phi) = e^{\beta (\phi - \phi_0)} \mu_0,$$

where $MA(\phi)$ is the Monge-Ampère measure of $\phi$, in the sense of Alexandrov (in particular, for $\phi$ smooth its density is the determinant of the Hessian appearing in the previous equation) and the condition [12] is assumed to hold in the sense of sub- gradients [27]. We may also, after a trivial scaling, assume that $P$ has unit Euclidean volume.

In the case when $\beta = 0$ (and, say, $\phi_0 = 0$) the corresponding Monge-Ampère equation, i.e. the equation

$$MA(\phi) = \mu_0,$$

where now $\mu_0$ is assumed to be a probability measure, plays a central role in the theory of optimal transport [35 39]. Under appropriate regularity assumptions on the measure $\mu_0$ a solution $\phi$ defines a map $T := \nabla \phi$ from $\mathbb{R}^n$ to $P$, which coincides with the so called optimal transport map, defined with respect to the the target measure $\lambda_P := 1_p dp$ (i.e. the normalized Lesbesgue measure on the convex body $P$) and the cost function $c(x,p) = -x \cdot p$. This means that it minimizes the corresponding total transport cost functional $C(T)$ over all maps $T$ transporting (i.e. pushing forward) $\mu_0$ to $\lambda_P$. The precise definitions are recalled in the appendix. For the moment we just recall that the cost functional $C$ is more generally defined on the space of all couplings $\Gamma$ (also called transference plans) between two given
probability measures $\mu$ and $\nu$, i.e $\Gamma$ is a measure on $\mathbb{R}^n \times \mathbb{R}^n$ whose push-forwards to the first and second factor are equal to $\mu$ and $\nu$, respectively and

$$C(\Gamma) := \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,p) \Gamma$$

Fixing $\nu = \lambda_P$ we will also write $C(\mu)$ for the corresponding optimal cost functional on the space of all probability measures on $\mathbb{R}^n$, i.e. $C(\mu)$ is the minimal cost to transport the measure $\mu$ to $\lambda_P$:

$$C(\mu) = \inf_{\Gamma} C(\Gamma)$$

where $\Gamma$ ranges over all couplings between $\mu$ and $\lambda_P$. It will also be important to consider the “weighted” cost functional $C_{\phi_0}(\mu)$ defined in terms of the cost function (1.4)

$$c_{\phi_0}(x,p) := -x \cdot p + \phi_0(x)$$

We will, in particular, show how to recover the solution $\phi$ of the equation 1.3 from the large $N$–limit of a certain random point process on $\mathbb{R}^n$ with $N$ particles, canonically determined by the given data $(\mu_0, \phi_0, P)$. In particular, this will, in the case $\beta = 0$, yield explicit explicit approximations for optimal transport maps. From the point of view of equilibrium statistical mechanics the parameter $\beta$ will play the role of the inverse temperature and the approach will involve the limiting zero temperature case (i.e. $\beta = \infty$) where, as explained below, the role of the equation 1.3 is played by a free boundary value problem for the Monge-Ampère equation, which can be equivalently described as a constrained convex envelope. As will be made clear below the results split into three different “phases” of increasing temperature: $\beta = \infty$, $\beta > 0$ and $\beta = 0$ (and we will also briefly comment on the negative temperature case $\beta < 0$ in section 7.3).

In the case $\beta = \infty$ we start with a weight function $\phi_0$ defined on a closed subset $X$ of $\mathbb{R}^n$ and the corresponding convex envelope $\phi_e$ is then defined as a point wise sup of convex functions $\phi$:

$$\phi_e(x) := \sup\{\phi(x) : \phi \leq \phi_0 \text{ on } X, \ \nabla \phi \in P\},$$

The Monge-Ampère measure of $\phi_e$ which is supported on $X$, will be denote by $\mu_e$. In fact, even if $X$ is non-compact it follows from the growth assumption on the weight $\phi_0$ that the support of $\mu_e$ is compact. When $\beta \to \infty$ the solutions of the corresponding Monge-Ampère equations 1.3 indeed converge, to the envelope $\phi_e$, where $X$ is the support of $\mu_0$, at least under an appropriate regularity assumption (see Prop 5.13). Interestingly, the Monge-Ampère measure $\mu_e$ also admits a natural interpretation in terms of the theory of optimal transport. Indeed, the measure $\mu_e$ minimizes, among all probability measures $\mu$ on $X$, the optimal transport cost $C_{\phi_0}(\mu)$ defined by the cost function in formula 1.4. Under suitable regularity assumptions (for example when $X = \mathbb{R}^n$ and is $\phi_0$ smooth) the map $T := \nabla \phi_e$ is the corresponding optimal transport map from the support of $\mu_e$ to the convex body $P$ (compare section 8.0.1).

Next, we turn to the definition of the corresponding random point processes, which will be defined as “$\beta$–deformations” of certain permanental random point processes, interpolating between a Poisson process for $\beta = 0$ and a permanental point process at $\beta = \infty$. First recall that the permanent of a rank $N$ matrix.
Let $A := (A_{i,j})$ be the real number defined by

$$\text{per}(A) := \sum_{\sigma \in S_N} \prod_{i=1}^N a_{i,\sigma(i)},$$

i.e. it is obtained from the definition of the determinant by removing the sign dependence on the permutation $\sigma$. Denote by $P_{\mathbb{Z}}$ the intersection of the convex body $\mathbb{Z}^n$ with the integer lattice $\mathbb{Z}^n$ and fix an auxiliary ordering $p_1,\ldots,p_N$ of the elements of $P_{\mathbb{Z}}$. Then the cost function determines a function $A(x_1,\ldots,x_N)$ on $(\mathbb{R}^n)^N$ with values in the space of $N$ times $N$ matrices, defined by

$$A_{ij}(x_1,\ldots,x_N) := e^{-c(x_i,p_j)}$$

and we will denote by $\text{Per}(x_1,\ldots,x_N)$ its permanent, which defines a real-valued function on $(\mathbb{R}^n)^N$, which is independent of the ordering of the lattice points $p_1,\ldots,p_N$. Indeed, we may write

$$\text{Per}(x_1,\ldots,x_N) = \sum_{\{p_1,\ldots,p_N\}} e^{x_1 \cdot p_1 + \cdots + x_N \cdot p_N},$$

where, for $(x_1,\ldots,x_N)$ fixed, the outer sum ranges over all possible $N!$ choices of $N$ different elements $p_1,\ldots,p_N$ in $P_{\mathbb{Z}}$. Given a weighted measure $(\mu_0,\phi_0)$ as above we then obtain a symmetric probability measure on $(\mathbb{R}^n)^N$, i.e. a random point process on $X$ with $N$ particles, by letting its density $\rho^{(N)}(x_1,\ldots,x_N)$, with respect to the product measure $\mu_0^{\otimes N}$, be proportional to the weighted permanent above, i.e.

$$\mu^{(N)} := \frac{1}{Z_N} \left( \text{Per}(x_1,\ldots,x_N)e^{-\phi_0(x_1)+\cdots+\phi(x_N)} \right) \mu_0^{\otimes N}$$

where $Z_N$ is the normalizing constant that will sometimes write as $Z_N[\phi_0]$ to indicate the dependence on $\phi_0$. This is an example of a permanental random point process on $X$ (see the survey for general properties of permanental point processes). Here we just recall that in the physics literature on many particle quantum systems such processes are used to represent a gas of free bosons and accordingly the name boson point processes is sometimes also used in the literature. In the present setting the corresponding bosons consist of $N$ “plane waves” $e^{ix \cdot k_i}$ with imaginary momenta (wave numbers) $k_i$ or more precisely: $k_i = -ip_i$ where $p_i$ ranges over the lattice points $P_{\mathbb{Z}}$ of $P$.

The empirical measure of the random point process introduced above, will be denoted by $\delta_N$. This is the random measure defined by

$$\delta_N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$$

We will consider the large $N$ limit which appears when $P$ is replaced by the sequence $kP$ of scaled convex bodies, for any positive integer $k$ (so that $N \sim k^n$ to the leading order) and the weight $\phi$ is replace by $k\phi$ and we will be concerned with the $\beta$-deformations of the permanental point process above, defined by the probability measure

$$\mu^{(N)}_{\beta_N} := \frac{1}{Z_{N,\beta_N}} \left( \text{Per}(x_1,\ldots,x_N)e^{-k(\phi_0(x_1)+\cdots+\phi(x_N))} \right)^{\beta_N/k} \mu_0^{\otimes N}$$

where the sequence $\beta_N$ is assumed to satisfy
\[ \lim_{N \to \infty} \beta_N = \beta \in [0, \infty) \]

(strictly speaking we should really write \( N = N_k \) to indicate the dependence on \( k \), but we have omitted the subscript \( k \) to simplify the notation). In particular, when \( \beta_N = k \) (so that \( \beta = \infty \)) we get a sequence of bona fide permanental point process. Note also that in the case when \( \beta_N \) is exactly equal to \( \beta \), for \( \beta \) finite, we could as well suppose that the weight function \( \phi_0 \) vanishes identically, by replacing \( \mu_0 \) with \( e^{-\beta \phi_0} \mu_0 \), but, in fact, it will be useful to separate the weight \( \phi_0 \) from the measure \( \mu_0 \).

**Theorem 1.1.** Given data \((\mu_0, \phi_0, P, \beta)\) as above the empirical measure \( \delta_N \) of the corresponding random point process on the set \( X \) (defined as the support of \( \mu_0 \)) converges in probability to the deterministic measure \( \mu_\beta \) defined by

\[ \mu_\beta = MA(\phi_\beta), \]

where \( \phi_\beta \) denotes the unique convex solution of equation 1.3 (satisfying 1.2) for \( \beta < \infty \) and for \( \beta = \infty \) the function \( \phi_\beta \) is the convex envelope 1.5. More precisely, the law of the empirical measure \( \delta_N \) (i.e. the probability measure \( (\delta_N, \mu_\beta(N)) \)) admits a large deviation principle (LDP) with rate \( N \beta \) and rate functional \( F_\beta \), where

\[ F_\beta(\mu) = C_{\phi_0}(\mu) + \frac{1}{\beta} D_{\mu_0}(\mu) - C_\beta, \]

where \( C_{\phi_0}(\mu) \) is the Monge-Kantorovich optimal cost functional corresponding to the cost function in formula 1.4, \( D_{\mu_0}(\mu) \) is the entropy of \( \mu \) relative to the background measure \( \mu_0 \) and \( C_\beta \) is the constant ensuring that the infimum of \( F_\beta \) is equal to 0.

The convergence in probability appearing the previous theorem is equivalent to the fact that the law of the empirical measure converges weakly to \( \delta_\beta \), the Dirac measure at \( \mu_\beta \). In turn, the LDP implies, since \( \mu_\beta \) is the unique minimizer of the rate functional \( F_\beta \), that the latter converge is exponential in a sense which may be loosely formulated as follows: denote by \( B_\delta(\mu) \) a ball of radius \( \delta \), centered at \( \mu \) in the space \( M_1(X) \) of all probability measure on \( X \), equipped with a metric defining the weak topology, then

\[ \text{Prob}(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} \in B_\delta(\mu)) \sim e^{-\beta N F_\beta(\mu)} \]

as \( N \to \infty \) and \( \delta \to 0 \) (see section 4.4 for the precise definition of the LDP). The Poisson case, i.e. when \( \beta = 0 \), is the content of Sanov’s classical theorem, which in turn is a generalization of Cramer’s theorem for random vectors in \( \mathbb{R}^n \) [24]. We also point of that the proof of Theorem 1.1 will give that \( C_\beta \) is equal to the following constant only depending on \((X, \phi_0, \beta)\)

\[ C_\beta := C(X, \phi_0, \beta) := \lim_{N \to \infty} -\frac{1}{N \beta_N} \log Z_{N, \beta_N}[\phi_0] = \inf_{\mu} \left( C_{\phi_0}(\mu) + \frac{1}{\beta} D_{\mu_0}(\mu) \right), \]

where, in the case \( \beta = \infty \) the infimum above is taken over all probability measures \( \mu \) supported on the support \( X \) of \( \mu_0 \).

Before commenting on the proof of the previous theorem we state some of its corollaries. First, from the convergence of the one-point correlation measures we obtain the following corollary, which provides a sequence of explicit approximate solutions to the real Monge-Ampère equations above.
Corollary 1.2. Let $\mu_0$ be a weighted measure and $P$ an $n$-dimensional convex body and fix $\beta > 0$. If $\mu_0 = \rho_0 dx$ for a strictly positive function $\rho_0$, then

$$
\phi_{\beta}^{(N)}(x) := \frac{1}{\beta} \log \int_{X^{N-1}} \frac{1}{Z_N} \left( \text{Per}(x, x_2, \ldots, x_N) \right)^{\beta/k} (e^{-\beta \rho_0(x_0)} \mu_0)^{\otimes N-1}
$$

satisfies (1.2) and converges locally uniformly, as $N \to \infty$, to the unique solution of the second boundary value problem with target $P$ for the Monge-Ampère equation (1.3).

In fact, in a similar way also obtain explicit approximations to the inhomogeneous Monge-Ampère equation (obtained by setting $\beta = 0$ in (1.1)). Formally, this is a consequence of the previous corollary in the limiting case $\beta = 0$, but the proof proceeds in a somewhat different manner.

Corollary 1.3. Let $\mu_0$ be a probability measure of the form $\mu_0 = \rho_0 1_X dx$ such that $X$ is the closure of a bounded domain whose boundary $\partial X$ is a null set for Lebesgue measure and assume that $\rho_0$ is bounded from below and above by positive constants on $X$. Then

$$
\phi^{(N)}(x) := \frac{1}{k} \int_{\mathbb{R}^{n(N-1)}} \log(\text{Per}(x, x_2, \ldots, x_N)) \rho_0(x_2) dx_2 \cdots \rho_0(x_N) dx_N - c_N,
$$

where $c_N$ is the normalizing constant ensuring that $\int_{\mathbb{R}^n} \phi^{(N)}(x) \rho_0 dx = 0$, converges, as $N \to \infty$, locally uniformly to the unique convex function $\phi$ solving the second boundary value problem with target $P$ for the Monge-Ampère equation $\text{MA}(\phi) = \mu_0$ with the normalization condition $\int_{\mathbb{R}^n} \phi \mu_0 = 0$. Moreover, $T^{(N)}(x) := \nabla \phi^{(N)}(x) :=$

$$
\frac{1}{k} \int_{\mathbb{R}^{n(N-1)} \sum_{\sigma \in S_N} P_{\sigma(1)} e^{P_{\sigma(1)} x_2 + P_{\sigma(2)} + \cdots + P_{\sigma(N)} x_N}} \left( x_2 \cdots \rho_0(x_N) dx_N \right)
$$

converges point-wise, in the interior of $X$, to the (Hölder continuous) optimal map $T$ for the Monge problem of transporting the probability measure $\mu_0$ on $X$ to the uniform probability measure $\lambda P$ on the target convex body $P$.

The existence of the optimal map $T$ in the previous corollary is due to Brenier [17] and the Hölder regularity was shown by Caffarelli [23]. As explained in section 6 a variant of the previous setting can also be considered which in particular applies to more general target measures $\nu$.

Coming back to Theorem 1.1 we point out that the key point in its proof is a rather general argument which, in a sense, reduces the problem to the case when $\beta = \infty$. We can then take $\beta_N = k$, so that the corresponding random point process is exactly permanental. In that case the theorem turns out to be a rather immediate consequence of the large deviation principle for determinantal point processes on polarized complex manifolds proved in [6], building on [4] [5]. More precisely, the latter result applies in the case when $P$ is a rational convex polytope, the point being that $P$ then determines a polarized toric variety $(X_P, L_P)$ with a projection map to $\mathbb{R}^n$ (see [2] and references therein). Anyway, we will give a direct purely "real" proof in the present setting for the case $\beta_N = k$ (the key ingredient is Prop 5.3). As for the reduction to the case $\beta = \infty$, it is inspired by some ideas from statistical mechanics and in particular mean field theory. In physical terms the idea of the argument may be explained as follows. Imagine that we know the macroscopic ground state (i.e. the state of zero energy $E$) of a system of a large number $N$ of particles in thermal equilibrium at zero temperature (i.e. at $\beta = \infty$). If we can rule
out any \textit{first order phase transitions} at zero-temperature (which essentially means that the macroscopic equilibrium state is unique), then increasing the temperature (i.e decreasing $\beta$) leads to a new macroscopic equilibrium state, minimizing the corresponding \textit{free energy} functional $E - S/\beta$, where $S$ is the physical entropy (i.e. $S = -D$ with our sign conventions). Increasing the temperature thus gives a transition from an ordered zero-temperature macroscopic state to a disordered positive temperature macroscopic state.

However, to apply this reasoning to the present situation we have to deal with $N$-particle interacting systems with rather general interactions ($N-$point Hamiltonians). One of the difficulties that we then have to confront is that the Hamiltonian in question is not a sum of two-point functions, as opposed to the more standard situation studied rigorously by Messer-Spohn \cite{34}. This is a reflection of the fact that the corresponding field equations \cite{1.1} are fully non-linear (i.e. non-linear in the derivative terms). More generally: the Hamiltonian is not given in the usual \textit{mean field} form, i.e. it is not of the functional form $H(\delta_N)$ for some $N-$independent functional $H$ on the space of probability measures on $X$ \cite{13}. Still, as explained in the next section we can obtain a rather general convergence result of mean field type for such interacting particle systems, that is hopefully of independent interest. As will be explained below this turns out to be related to previous work of Ellis-Have-Turkington \cite{25}. The result applies in particular to the setting above where the corresponding Hamiltonian may (in the non-weighted case) be written as $H^{(N)}(x_1, \ldots, x_N) = -\frac{1}{k} \log \text{Per}(x_1, \ldots, x_N)$, i.e.

\begin{equation}
H^{(N)}(x_1, \ldots, x_N) = -\frac{1}{k} \log \sum_{\sigma \in S_N} e^{-kNC(\sigma)} , \quad C(\sigma) := -\frac{1}{k}(x_1 \cdot p_{\sigma(1)} + \cdots + x_N \cdot p_{\sigma(N)})/N
\end{equation}

As a side remark we note that this form of writing the Hamiltonian gives a simple heuristic interpretation of the large deviation result for $\beta_N = k$; the result essentially says that we may, when $N \to \infty$, replace the whole sum over all permutations with the contribution from the permutation with minimal cost $C(\sigma)$ (compare Remark \cite{5.9} and the relation between optimal transport and its discrete version described in the appendix). But the actual proof proceeds in a different way, based on a duality argument, where the cost functional $C(\mu)$ arises as the Legendre transform of another functional on the space $C_b(X)$ of all bounded continuous functions on $X$.

As explained below, for a fixed $\beta$, the Monge-Ampère measure $\mu_\beta(= MA(\phi_\beta))$, where $\phi_\beta$ is the solution of equation \cite{1.3} arises as the minimizer of the corresponding free energy functional. From this point of view it is interesting to study the behaviour of the measures $\mu_\beta$, describing the equilibrium states at inverse temperature $\beta$, as $\beta$ varies. For example, when the fixed background measure $\mu_0$ is equal (or comparable) to the usual Euclidean measure $dx$ the support of $\mu_\beta$ is all of $\mathbb{R}^n$. However, the limiting measure obtained when $\beta$ increases to infinity is always supported on a \textit{compact} set, whose “boundary” appears as the free boundary for a Monge-Ampère equation, as explained above. This shows that the the phase transition at $\beta = 0$ referred to above is, in the present setting, reminiscent of a liquid-gas phase transition.

Before continuing we also point out that the main new technical difficulty which appears in the complex geometric setting out-lined in \cite{10}, where the role of the
permanent above is played by a Vandermonde type determinant is that the corresponding Hamiltonian $H^N(z_1, ..., z_N)$ is singular (for example when different points merge). See [13, 30] for the case of dimension $n = 1$, or more precisely the case when $X$ is a domain in the complex plane $\mathbb{C}$. A major simplifying feature which appears in the case $n = 1$ is that the Vandermonde determinant factorizes completely and the point process in question is hence a Coulomb gas, i.e. $H^N(z_1, ..., z_N)$ is then proportional to a sum of two-point functions of the form $\log |z_i - z_j|$ (compare the end of section 7.2).

1.2. Interacting particle system in thermal equilibrium. Let $X$ be a closed set in $\mathbb{R}^n$ and assume, for simplicity, that $X$ is compact (generalizations to the non-compact case of $\mathbb{R}^n$ will be considered in section 5). Fix a probability measure $\mu_0$ supported on $X$. For a fixed positive integer $N$ (representing the number of particles) we assume given an $N$–particle Hamiltonian $H^N$, i.e. a continuous function on the $N$–fold product $X^N$ which is symmetric, i.e. invariant under the action of the permutation group $S_N$. The corresponding Gibbs measure is the probability measure on $X^N$ defined by

$$
\mu_{\beta}^{(N)} := e^{-\beta N H^N} \mu_0^{\otimes N} / Z_N, \beta_N,
$$

where we have also fixed a positive number $\beta_N$ (the inverse temperature) and where $Z_N$ is the normalizing constant (partition function), i.e.

$$
Z_{N, \beta_N} := \int_{X^N} e^{-\beta N H^N} \mu_0^{\otimes N}.
$$

We will also assume that $H^N$ is uniformly Lipschitz continuous (in fact, equicontinuous will be enough; compare section 2.1). In the case when $H^N$ is differentiable this thus simply means that there is a constant $L$ independent of $N$ such that

$$
\sup_X \left| \partial_u H^N(x_1, ..., x_N) \right| \leq L
$$

Given a continuous function $u$ on $X$ we denote by $Z_{N, \beta_N}[u]$ the “tilted” partition function obtained by replacing $H^N$ with $H^N + u$, where $u(x_1, ..., x_N) := \sum u(x_i)$. In other words $Z_{N, \beta_N}[u]$ is the (scaled) Laplace transform of the law of the empirical measure $\delta_N$ defined by the Gibbs measure associated to $H_N$.

**Theorem 1.4.** Let $H^N$ an $N$–particle Hamiltonian as above. Assume that there exists a sequence $\beta_N$ of positive real numbers tending to infinity such that $-\beta_N \log Z_{N, \beta_N}[u]$ converges, when $N \to \infty$, to a functional $F(u)$ which is Gateaux differentiable on $C^0(X)$. Then, for any fixed $\beta > 0$, the law of the empirical measure $\delta_N$ for the random point process on $X$ defined by the corresponding Gibbs measure $\mu_{\beta}^{(N)}$ satisfies a LDP with speed $\beta N$ and good rate functional

$$
F_{\beta}(\mu) = E(\mu) + \frac{1}{\beta} D_{\mu_0}(\mu) - C_{\beta}
$$

where $C_{\beta}$ is the normalizing constant and where the functional $E(\mu)$, which only depends on the support $X$ of $\mu_0$, denotes the Legendre transform of $F(u)$ (formula 2.5) and $D_{\mu_0}(\mu)$ is the entropy of $\mu$ relative to $\mu_0$. In particular, $\delta_N$ converges in probability to the unique minimizer $\mu_{\beta}$ of $F_{\beta}$.
Moreover, we will show that the minimizer $\mu_\beta$ can be written as $\mu_\beta = dF|_u$ where $u_\beta$ is a continuous function on $X$ solving the following equation of mean field type:

$$dF|_u = e^{\beta u} \mu_0,$$

where $dF|_u$ denotes the measure defined by the Gateaux differential of $F$ at $u$ (see Theorem 3.6). We will also obtain a canonical sequence $u_N$ of functions uniformly converging to $u$, which are the unique solutions of certain “finite $N$” approximations to the mean field type equation above and which arise as the limiting fixed point of certain iterations (see section 3.2). The construction of $u_N$ is inspired by Donaldson’s notion of balanced metrics, introduced in the setting of Kähler-Einstein metrics [19]) and is hopefully of independent interest.

Theorem 1.4 generalizes the seminal result of Messer-Spohn [34] and is closely related to a previous result of Ellis-Haven-Turkington [25] (compare section 4.6). In the statement of the theorem we have made the rather strong assumptions of (a) compactness of $X$ and (b) equicontinuity of $H(N)$. In fact, under these assumptions a more direct proof can be given by essentially reducing the problem to the classical LDP in the non-interacting case (i.e. $H(N) = 0$ or equivalently $\beta = 0$) which is the content of Sanov’s theorem (see section 4.6 where a comparison with the work [25] is also made). But the main point here is to give a flexible proof that can be adapted to more general situations where the assumptions (a) and (b) may not be satisfied. For example, in the setting of Theorem 1.1 above the the compactness assumption (a) does not hold and in the complex geometric setting refereed to above the assumption (b) is not satisfied, since $H(N)$ is even singular (compare section 7.2).

The key step in the proof of the previous theorem is to establish the convergence of the limiting mean energy towards the Legendre transform of $F$:

$$\lim_{N \to \infty} \int_{X^N} \frac{H(N)}{N} \mu^{\otimes N} = F^*(\mu)$$

(Theorem 2.1). Then applying the finite dimensional Gibbs variational principle and some variational arguments allows us to compute the limit of the corresponding partition functions, “tilted by $u$”. Differentiating at $u = 0$ this is implies in particular the convergence of the one-point correlation measures towards the minimizer of the free energy functional. Finally, a suitable application of the (generalized) Gärtner-Ellis theorem gives the LDP principle in the theorem. This approach is inspired by the approach of Messer-Spohn [34]. But the key new observation here is that (even if the assumptions (a) and (b) are not satisfied) the desired LDP holds as long as the convergence in formula 1.9 holds (in fact, the lower bounds always holds). More precisely, it is enough if the convergence holds for all $\mu$ in the subspace $\{F_\beta < \infty\}$ as long as the latter subspace is dense.

Another point of the approach developed here, as compared to [34], is to separate the case $\beta > 0$ from the more subtle case $\beta < 0$, by bypassing the use of the Hewitt-Savage decomposition theorem used in [34]. In fact, if one uses the latter decomposition theorem then one obtains, in the spirit of [34], a generalization of Theorem 1.4 to possibly negative $\beta$, saying that, after passing to a subsequence, the law of the empirical measures converge weakly to a probability measure on the space of all probability measures which is concentrated on the minimizers of the functional $F_\beta$. The point is that in the case $\beta > 0$ the functional $F_\beta$ is strictly convex and
hence admits a unique minimizer, as in Theorem 1.4. However, extensions to the non-compact setting of $\mathbb{R}^n$, in the case $\beta < 0$, require refined growth estimates and we thus leave this case for a separate paper where applications to Kähler-Einstein geometry on toric varieties will be given [11]. For a brief outline of the case $\beta < 0$ in the Monge-Ampère setting in $\mathbb{R}^n$ and the relations to the existence problem for Kähler-Einstein metrics on complex algebraic varieties and phase transitions, see section 7.3.

Outline of the paper. After having introduced the precise assumptions on the $N$–particle Hamiltonian in section 2 we prove the existence of the limiting mean energy for measures with continuous potentials. This assumption is automatically satisfied in the Monge-Ampère setting and the existence in the general case is proved in section 3 by establishing an approximation result of independent interest, inspired by Donaldson’s notion of balanced metrics. Then in section 4 we go on to establish the LDP in Theorem 1.4 and we also give an alternative more direct proof of the latter theorem. In section 5 we introduce the Monge-Ampère setting leading up to a proof of Theorem 1.1. In section 6 a variant of the previous setting is considered which in particular applies to rather general target measures. A comparison with the complex geometric setting and an outlook on further developments is given in section 7. The paper is concluded with an appendix giving some background on optimal transport and establishing the (essentially well-known) comparison principle in the Monge-Ampère setting.

Acknowledgments. Thanks to Ofer Zeitouni for helpful comments on a first draft of this paper and Bo’az Klartag for stimulating discussions. As pointed out in the introduction, this paper and in particular the connections to the theory of optimal transport arose as a “spin-off effect” of some of my work in complex geometry and pluripotential theory. Lacking background in optimal transport theory I apologize for any omission of accrediting prior results properly. This research was supported by grants from the European Research Council (ERC starting grant 307529) and the Swedish Research Council.

2. Existence of the limiting mean energy

2.1. Setup and assumptions on the Hamiltonian $H^{(N)}$. Let $(X,d,\mu_0)$ be a compact metric space with a probability (Borel) measure $\mu_0$, where $X$ is the support of $\mu_0$. The typical situation is when $X$ is embedded in a Riemannian manifold and $d$ is induced from the Riemannian metric.

Let $H^{(N)}$ be an $N$–particle Hamiltonian, i.e. continuous function on the $N$–fold product $X^N$ which is symmetric, i.e. invariant under the action of the permutation group $S_N$. We denote by $d_{X^N}$ the following induced distance function on $X^N$:

$$d_{X^N}((x_1, \ldots, x_N), (y_1, \ldots, y_N)) := \frac{1}{N} \sum_{i=1}^{N} d_X(x_i, y_i)$$

We will assume that $H^{(N)}/N$ is equicontinuous, i.e. given any $\epsilon > 0$ there exists a $\delta > 0$ such that $|H^{(N)}(x) - H^{(N)}(y)|/N \leq \epsilon$ if $d_{X^N}(x, y) \leq \delta$. It will often be convenient to make the stronger assumption that $H^{(N)}/N$ be uniformly Lipschitz continuous in the sense that, there exists a constant $L$ such that

$$|H^{(N)}(x) - H^{(N)}(y)|/N \leq L d_{X^N}(x, y)$$

(2.1)
In particular, the latter property holds if $H^{(N)}(x, x_2, \ldots, x_N)$ is Lipschitz continuous on $X$ with Lipschitz constant $L$, for any fixed $(x_2, \ldots, x_N) \in X^{N-1}$.

We will also assume that $H^{(N)}$ has the following limiting properties:

- The following limit exists for any $u \in C^0(X)$:

$$F(u) := \lim_{N \to \infty} \inf_{X^{N-1}} \frac{1}{N} H^{(N)}(u),$$

where we have used the notation $u(x_1, \ldots, x_N) := \sum_i u(x_i)$.

- The functional $F$ is Gateaux differentiable on $C^0(X)$ (i.e. differentiable along all affine lines in $C^0(X)$)

It will turn out that, from the statistical mechanical point of view, the first part is equivalent to the “existence of the free energy” (compare Lemma 2.6).

We will write $X(N) := X/S_N$, where $S_N$ is the permutation group with $N$ elements and equip it with the induced quotient distance function

$$d_{X(N)}((x_1, \ldots, x_N), (y_1, \ldots, y_N)) := \inf_{\sigma \in S_N} \frac{1}{N} \sum_{i=1}^{N} d_X(x_i, y_{\sigma(i)})$$

and denote by $M_1(X)$ the space of all probability measure $\mu$ on $X$. Similarly, we denote by $M_1(X(N))$ the $S_N$-invariant subspace of $M_1(X^N)$. There is a standard embedding

$$X(N) \hookrightarrow M_1(X), \ (x_1, \ldots, x_N) \mapsto \delta_N := \frac{1}{N} \sum \delta_{x_i}$$

where we will call $\delta_N$ the empirical measure. We equip the space $M_1(X)$ with the Wasserstein 1-metric $d_W$ determined by $(X, d)$. Then it is well-known that the embedding above becomes an isometry (see the end of section 8.0.1), i.e.

$$d_W(\delta_N(x), \delta_N(x')) = d_{X(N)}(x, x')$$

When $X$ is compact (which we assume in this section) the topology on $M_1(X)$ induced by $d_W$ coincides with the usual topology defining weak convergence of measures.

### 2.2. Existence of the limiting mean energy

Given an element $\mu$ in the space $M_1(X)$ of probability measures on $X$ we define its thermodynamic energy $E(\mu)$ as the Legendre transform of $F$ at $\mu$ (with a somewhat non-standard sign convention):

$$E(\mu) := \sup_{u \in C^0(X)} (F(u) - \int_X u \mu) := \sup_{u \in C^0(X)} F(\mu)$$

The key point is the following result which identifies $E$ with the limiting mean energy (this terminology will be taken up in section 4) under the assumption that $F$ be Gateaux differentiable.

**Theorem 2.1.** Let $H^{(N)}$ be an $N$-particle Hamiltonian satisfying the properties in section 2.1 then

$$\lim_{N \to \infty} \frac{1}{N} \int_{X^N} H^{(N)} \mu^N = E(\mu)$$

for any probability measure $\mu$ on $X$. 

11
2.3. The proof of Theorem 2.1 Let us start by fixing $\mu$ and $u \in C^0(X)$ and rewriting

$$
\frac{1}{N} \int_X H^{(N)} \mu^{\otimes N} = \left( \inf_{X^N} \frac{H^{(N)} + u}{N} - \int_X u \mu \right) + I_N[\mu, u],
$$

where

$$
I_N[\mu, u] := \int_X \left( \frac{H^{(N)} + u}{N} - \inf_{X^N} \frac{H^{(N)} + u}{N} \right) \mu^{\otimes N}.
$$

Since, trivially, $I_N[\mu, u] \geq 0$ it follows that the lower bound in the theorem to be proved always holds:

$$
E(\mu) := \sup_{u \in C^0(X)} \left( \inf_{X^N} \frac{H^{(N)} + u}{N} - \int_X u \mu \right) \leq \liminf_{N \to \infty} \frac{1}{N} \int_X H^{(N)} \mu^{\otimes N},
$$

To handle the upper bound we will use the assumed differentiability of the functional $F(u)$. Let us start by introducing the following terminology:

**Definition 2.2.** A function $u_\mu$ on $X$ is said to be a potential of the measure $\mu$ if $u_\mu$ is a maximizer of the functional whose sup defines $E(\mu)$, i.e. if

$$
E(\mu) := F(u_\mu) - \int_X u_\mu \mu.
$$

Since $F$ is concave and also assumed Gateaux differentiable this equivalently means that

$$
\mu = dF[\mu^{\ast}].
$$

The key upper bound that we will prove may be formulated as the following

**Proposition 2.3.** Assume that the functional $F$ is Gateaux differentiable. Then, for any probability measure $\mu$ admitting a continuous potential we have

$$
\limsup_{N \to \infty} \frac{1}{N} \int_X H^{(N)} \mu^{\otimes N} \leq E(\mu).
$$

In order to prove the previous proposition we start with the following simple but very useful lemma (which was used in the similar context of Fekete points in [5]).

**Lemma 2.4.** Fix $u^{\ast} \in C^0(X)$ and assume that $x^{\ast(N)} \in X^N$ is a minimizer of the function $(H^{(N)} + u^{\ast})/N$ on $X^N$. If the corresponding large $N-$ limit $F(u)$ exists for all $u \in C^0(X)$ and $F$ is Gateaux differentiable at $u^{\ast}$, then $\delta_N(x^{\ast(N)})$ converges weakly towards $\mu^{\ast} := dF[\mu^{\ast}]$.

**Proof.** Fix $v \in C^0(X)$ and a real number $t$. Let $f_N(t) := \frac{1}{N}(H^{(N)} + u + tv)(x^{\ast(N)})$ and $f(t) := F(u + tv)$. By assumption $\lim_{N \to \infty} f_N(0) = f(0)$ and $\liminf_{N \to \infty} f_N(t) \geq f(t)$. Note that $f$ is a concave function in $t$ (since it is defined as an inf of affine functions) and $f_N(t)$ is affine in $t$. But then it follows from the differentiability of $f$ at $t = 0$ that $\lim_{N \to \infty} df_N(t)/dt|_{t=0} = df(t)/dt|_{t=0}$, i.e. that

$$
\lim_{N \to \infty} \left< \delta_N(x^{\ast(N)}), v \right> = \left< df[\mu^{\ast}], v \right>,
$$

which thus concludes the proof of the lemma.

Let us also recall the following weak form of Sanov’s theorem:
Lemma 2.5. For any given \( \mu \in \mathcal{M}_1(X) \) we have that \((\delta_N)_* \mu^{\otimes N} \) converges to \( \delta_\mu \) in the space \( \mathcal{M}_1(X) \) equipped with the weak topology.

Proof. In fact, this is an immediate consequence of the following useful result: if \( \mu_N \in \mathcal{M}_N(X^N) \) then \((\delta_N)_* \mu_N \to \delta_\mu \) iff the corresponding \( j \)-point correlation measures \( \int_{X^{N-j}} \mu_N \) converge weakly to \( \mu^{\otimes j} \) (see Prop. 2.2 in [37]). \( \square \)

2.3.1. The proof of Prop 2.3. By Theorem 3.5 below any probability measure \( \mu \) admits a continuous potential \( u_\mu \). Given the measure \( \mu \in \mathcal{M}_1(X) \) we may thus assume that the potential \( u_\mu \) is continuous. Since \( X \) is compact this means that for any given \( \epsilon > 0 \) there exists \( \delta > 0 \) such that \( d_X(x, x') < 2\delta \) implies that \( |u_\mu(x) - u_\mu(x')| \leq \epsilon \). Moreover, by the assumption on \( H(N)/N \) the corresponding property of \( H(N)/N \) holds uniformly in \( N \) on \( X^N \). In particular,

\[
(2.9) \quad d_{X^N}(x, x') \leq 2\delta \Rightarrow |\left( \frac{H(N)}{N} + u_\mu \right)(x) - \left( \frac{H(N)}{N} + u_\mu \right)(x')| \leq 2\epsilon
\]

We denote by \( B_\delta(\mu) \) the inverse image in \( X^N \) of a ball of radius \( \delta \) centered at \( \mu \) in \( (\mathcal{M}_1(X), d_W) \), under the map \( \delta_N \). Let us first note that

\[
(2.10) \quad \lim_{N \to \infty} \int_{X^N \setminus B_\delta(\mu)} \frac{H(N)}{N} + u_\mu \mu^{\otimes N} = 0
\]

Indeed, since \( \frac{H(N) + u_\mu}{N} \) is uniformly bounded this follows immediately from the previous lemma. Let now \( x_\mu(N) \) be a sequence as in Lemma 2.4 associated to \( u = u_\mu \). By the lemma we have that \( \delta_N(x_\mu(N)) \) converges weakly to \( \mu \). Hence, taking \( N \) sufficiently large \( (N \geq N_\delta) \) we may as well assume that \( x_\mu(N) \in B_\delta(\mu) \). As a consequence, if \( x \) is any given point in \( B_\delta(\mu) \) then, by the isometry property 2.4 \( d_{X^N}(x, x_\mu(N)) \leq 2\delta \). But then 2.9 gives

\[
(2.11) \quad \int_{B_\delta(\mu)} \left( \frac{H(N) + u_\mu}{N} \right) - \inf_{X^N} \left( \frac{H(N) + u_\mu}{N} \right) \leq 2\epsilon
\]

Combining 2.10 and 2.11 we finally deduce that \( \limsup_{N \to \infty} I_N[\mu, u_\mu] \leq \epsilon \), for any \( \epsilon > 0 \), which thus concludes the proof of Prop 2.3.

2.4. The mean energy as the Legendre transform of the limiting free energy. It will be very useful to obtain an alternative integral expression for the functional \( \mathcal{F}(u) \). To this end we fix a probability measure \( \mu_0 \) with support \( X \) and a sequence \( \beta_N \to \infty \) of positive numbers and set

\[
(2.12) \quad \mathcal{F}(N)(u) := -\frac{1}{\beta_N N} \log \int_{X^N} e^{-\beta_N N (H(N)/N + u)} \mu_0^{\otimes N},
\]

where we have omitted the explicit dependence on \( \mu_0 \) in the notation (in terms of the notation introduced in section we thus have the \( \mathcal{F}(N) \) may be written as \( \mathcal{F}(N)(u) = -\frac{1}{\beta_N} \log Z_{N, \beta_N}[u] \).

Lemma 2.6. Assume that \( H(N)/N \) is equicontinuous (as defined in section 2.1) and that \( u \) is continuous on \( X \). Then, as \( N \to \infty \),

\[
\inf_{X^N} \frac{1}{N} (H(N) + u) = \mathcal{F}(N)(u) + o(1)
\]

for any sequence \( \beta_N \to \infty \).
Proof. Setting $G^{(N)} := -(H^{(N)} + u)/N$, then it will be enough to show that for any $\epsilon > 0$ there exists a constant $C_\epsilon$ such that

$$\frac{1}{\beta N N} \log \int_{X^N} e^{\beta N G^{(N)} \mu_0 \otimes N} \geq \sup_{X^N} G^{(N)} - C_\epsilon / \beta N - \epsilon,$$

To this end we denote by $x^{(N)}_*$ a configuration where the sup in the rhs above is attained and restrict the integration to a polydisc $\Delta_\delta(x^{(N)}_*)$ of radius $\delta$ centered at $x^{(N)}_*$. Thus

$$\frac{1}{\beta N N} \log \int_{X^N} e^{\beta N (G^{(N)} \mu_0 \otimes N)} \geq \sup_{X^N} G^{(N)} + \frac{1}{\beta N N} \log \int_{\Delta_\delta(x^{(N)}_*)} e^{\beta N (G^{(N)} - G^{(N)}(x^{(N)}_*))}$$

Now, by the uniform continuity assumption, given $\epsilon > 0$ we can take $\delta > 0$ such that

$$\frac{1}{\beta N N} \log \int_{\Delta_\delta(x^{(N)}_*)} e^{\beta N (G^{(N)} - G^{(N)}(x^{(N)}_*))} \geq -\epsilon + \frac{1}{\beta N N} \log \int_{\Delta_\delta(x^{(N)}_*)} \mu_0 \otimes N$$

Since, by assumption, the “coordinates” of $x^{(N)}_*$ are contained in the support of $\mu_0$ the last term above may, for any $\delta > 0$ be estimated from below by $\frac{1}{\beta N N} \log(C_\delta)^N$ which thus concludes the proof (a similar argument is used in the proof of Lemma 5.2).

By the previous Lemma the first point in the assumptions about the limiting properties of $H^{(N)}$ (in section 2.1) is thus equivalent to the existence of the limit

$$\mathcal{F}(u) := \lim_{N \to \infty} \mathcal{F}^{(N)}(u)$$

and in particular the limit only depends on the support $X$ of $\mu_0$ and not on $\mu_0$ itself.

3. Generalized mean field equations and balanced functions

In this section we will, among other thing, prove that any probability measure admits a continuous potential, which was used in the proof of Theorem 2.1. The theorem will be deduced from a general result of independent interest (Theorem 3.3) which will allow us to write $u_\mu$ as a $C^0$–limit of equicontinuous functions $u_N$ maximizing the functional

$$(3.1) \quad \mathcal{F}^{(N)}(u) := \mathcal{F}(u) - \int u \mu$$

In fact, this argument will give a more precise result saying that $u_\mu$ is in a certain subspace $\mathcal{P}(X)$ of $C^0(X)$ which, in the setting of the real Monge-Ampère operator may be identified with the space of all convex functions whose gradient image is contained in the given target convex body $P$. In the latter setting the functions $u_N$ essentially coincide with Donaldson’s $\mu$–balanced metrics (see [19] [7]).

This approach will also lead to a natural setting for formulating global equations, that we will refer to as generalized mean field equations, which generalize the real Monge-Ampère equation [3,3].
3.1. Setup. To simplify the exposition of the proof it will be convenient to assume that $H^{(N)}$ is uniformly Lipschitz continuous with Lipschitz constant $L$, but the proof under the more general assumption of uniform continuity is essentially the same. Let $L(X)$ be the space of all Lipschitz continuous functions on $X$ with Lipschitz constant $L$. We let $\mathcal{P}_N(X)$ be the subspace of $L(X)$ of all functions $u$ such that there exists a finite measure $\nu$ on $X^N$ such that

$$u(x) = \frac{1}{\beta_N} \log \int_{y \in X^{N-1}} e^{-\beta_N H^{(N)}(x,y)} \nu$$

Next, we define $\mathcal{P}(X)$ as the closure in $C^0(X)$ of the union of all spaces $\mathcal{P}_N(X)$ as $N$ ranges over all positive integers. By construction we thus have

$$\mathcal{P}_N(X) \subset \mathcal{P}(X) \subset L(X) \subset C^0(X)$$

3.2. Balanced functions and their large $N$–limit. Let $\pi_N$ be the operator

$$\pi_N : C^0(X) \rightarrow \mathcal{P}_N$$

defined by

$$\pi_N(u)(x) := \frac{1}{\beta_N} \log \frac{1}{Z_N[u]} \int_{y \in X^{N-1}} e^{-\beta_N H^{(N)}(x,y)+u(y)} \otimes (N-1),$$

where, as usual, $Z_N[u] := \int_{X^N} e^{-\beta_N H^{(N)}(x)+u(x)} \otimes \mu_0^{\otimes N}$. The definition of $\pi_N(u)$ is made so that $e^{\beta_N (\pi_N(u)-u)} \mu_0$ is a probability measure (in fact, $e^{\beta_N (\pi_N(u)-u)} \mu_0$ coincides with the one-point correlation measure $\mu^{(N)}_1$ of the corresponding Gibbs measure; compare section 4.1).

**Definition 3.1.** A function $u_N$ on $X$ is said to be balanced with respect to $(\mu_0, \beta_N)$ if $\pi_N(u_N) = u_N$.

**Proposition 3.2.** Given a probability measure $\mu_0$ on $X$ there exists, for any integer $N$, a function $u_N \in \mathcal{P}_N(X)$ which is balanced with respect to $(\mu_0, \beta_N)$. Moreover, $u_N$ maximizes the functional $\mathcal{F}_{\mu_0}^{(N)}(u)$ on $C^0(X)$ and is uniquely determined mod $\mathbb{R}$.

**Proof:** First observe that, by definition, $\pi_N(u + c) = \pi_N(u) + c$ for any constant $c$ and hence $\pi_N$ descends to a map on $L(X)/\mathbb{R}$. By the Arzelà-Ascoli theorem the latter space is a compact subspace of the quotient Banach space $C^0(X)/\mathbb{R}$ (where $C^0(X)$ is equipped with the usual $C_0$–norm $\|u\| := \sup_X |u|$). The existence of $u_N$ now follows from the Schauder fixed point theorem applied to the continuous operator $\pi_N$ acting on the compact convex space $L(X)/\mathbb{R}$ (the continuity is an immediate consequence of the explicit expression (3.3)). Note that, by the mapping property (3.2) $u_N$, is in fact contained in the subspace $\mathcal{P}_N(X)$ of $L(X)$. To conclude the proof of the proposition we observe that a direct calculation reveals that the differential of $\mathcal{F}^{(N)}$ is given by the following formula:

$$d(\mathcal{F}^{(N)})(u) = e^{\beta_N (\pi_N(u)-u)} \mu_0$$

Hence, if $\pi_N(u_N) = u_N$ then $d(\mathcal{F}^{(N)})(u_N) = \mu_0$ which equivalently means that $u_N$ is a critical point of the functional $\mathcal{F}_{\mu_0}^{(N)}$. The maximization property then follows directly from the fact that $\mathcal{F}_{\mu_0}^{(N)}$ is concave (since $\mathcal{F}^{(N)}$ is, as follows directly from the concavity of log). Finally, since $\mathcal{F}_{\mu_0}^{(N)}$ is in fact strictly concave mod $\mathbb{R}$ (by the
strict concavity of log) the uniqueness of a critical point modulo additive constants follows. \hfill \square

**Theorem 3.3.** Let \( \mu_0 \) be a probability measure with compact support \( X \) and \( H^{(N)} \) an \( N \)-particle Hamiltonian satisfying the assumptions in section 2.1. Let \( \mu \) be another probability measure on \( X \) with the same support \( X \) as \( \mu_0 \). Then there exists (after perhaps passing to a subsequence) a sequence of functions \( u_N \) in \( P_N(X) \) which are balanced with respect to \( (\mu, \beta_N) \) and such that \( u_N \to u_\mu \) in \( C^0(X) \), where \( u_\mu \) is a potential for \( \mu \), i.e.

\[
dF|_{u_\mu} = \mu
\]

**Proof.** Let us first prove the result for \( \mu = \mu_0 \). Let \( u_N \) be a sequence of functions which are balanced wrt \( (\mu_0, \beta_N) \). Since \( \pi_N(u+c) = \pi_N(u) + c \) we may as well assume that \( u_N \) is normalized in the sense that \( u_N(x_0) = 0 \) at some fixed point \( x_0 \). By the Arzelà-Ascoli theorem there exists, after perhaps passing to a subsequence, a function \( u_\infty \in L(X) \) such that \( u_N \to u_\infty \) in \( C^0 \)-norm. Let us next show the following

\[
(3.4) \quad \text{Claim: } u_\infty \text{ maximizes } F_{\mu_0} \text{ on } C^0(X)
\]

First, by the previous proposition, \( u_N \) maximizes \( F^{(N)}_{\mu} \) and hence if \( u \) is a fixed element in \( C^0(X) \) we get

\[
F_{\mu_0}(u) := \lim_{N \to \infty} F^{(N)}_{\mu_0}(u) \leq F^{(N)}_{\mu_0}(u_N)
\]

Next, by construction \( u_N \leq u_\infty + \delta_N \), where \( \delta_N \) is a sequence of positive numbers tending to zero and hence

\[
F^{(N)}_{\mu_0}(u_N) \leq F^{(N)}_{\mu_0}(u_\infty + \delta_N) = F^{(N)}_{\mu_0}(u_\infty),
\]

where, by definition, the rhs above converges to \( F_{\mu_0}(u_\infty) \) as \( N \to \infty \). This proves the claim above. Since \( F \) is Gateaux differentiable it follows that the differential \( dF_{\mu_0} \) vanishes at \( u_\mu := u_\infty \), which translates to the equation in the theorem. \hfill \square

Before turning to the proof of 3.3 we state the following corollary of the previous theorem (or rather its proof), which gives yet another formula for the energy functional \( E \) defined by 2.6.

**Corollary 3.4.** Let \( \mu \) be a probability measure with support \( X \) and \( \beta_N \) a sequence tending to infinity. Denote by \( u_N \in \mathcal{P}_N(X) \) a sequence of functions which are balanced wrt \( (\mu, \beta_N) \). Then

\[
E(\mu) = \lim_{N \to \infty} \sup_{u \in \mathcal{P}_N(X)} \left( F^{(N)}_{\mu}(u) - \int u \mu \right) = \lim_{N \to \infty} F^{(N)}_{\mu}(u_N) - \int u_N \mu
\]

3.3. **Existence of continuous potentials.** We can now prove the existence of continuous potentials:

**Theorem 3.5.** Let \( H^{(N)} \) be an \( N \)-particle Hamiltonian satisfying the assumptions in section 2.1 and denote by \( F(u) \) the corresponding limiting functional (formula 2.2). Then any probability measure \( \mu \) on \( X \) admits a continuous potential \( u_\mu \). Equivalently, \( u_\mu \) solves the equation

\[
dF|_{u_\mu} = \mu
\]

Moreover, if \( H^{(N)} \) admits a uniform Lipschitz constant \( L \), then so does \( u_\mu \).
Proof. In the case when \( \mu \) has the same support as \( \mu_0 \) (i.e. the set \( X \)) the result follows immediately from the previous theorem (since the limiting functional \( \mathcal{F} \) only depends on the support of \( X \), by Lemma 2.4). In the general case we instead first apply Theorem 3.3 to \( \mu_* := (1 - \epsilon) \mu + \epsilon \mu_0 \) and obtain (normalized) potentials \( u_\epsilon \) of \( \mu_* \) in \( L(X) \). After passing to a subsequence we may assume that \( u_\epsilon \to u \) in \( C^0 \)-norm. The proof is now concluded by noting that \( u \) maximizes the functional \( \mathcal{F}_\mu \), as proved by a slight modification of the proof of the claim appearing in the proof of Theorem 3.3. \qed

3.4. The generalized mean field equations. In this section we will show that the minimizer of the free energy functional \( F_\beta \) can be obtained from the solutions of the equation (3.8) appearing in the introduction of the paper, which can be seen as a generalization of the real Monge-Ampère equation (3.3). In fact, this part of the argument was carried out in the more general setting of complex Monge-Ampère equations in [8], which is analytically considerably more involved.

We first define

\[
G_{\mu_0, \beta}(u) := \mathcal{F}(u) - \frac{1}{\beta} \log \int e^{\beta u} \mu_0
\]

whose critical point equation is

\[
d\mathcal{F}_{\mu_0} = \frac{e^{\beta u} \mu_0}{\int e^{\beta u} \mu_0}
\]

By the Gateaux differentiability and concavity of \( \mathcal{F}(u) \) we have that \( u \) satisfies the previous equation iff \( u \) maximizes \( G_{\mu_0, \beta} \). The equation above is invariant under the additive action of \( \mathbb{R} \) on \( C^0(X) \) and may hence be formulated as an equation on \( C^0(X)/\mathbb{R} \) which in turn is equivalent to the following equation on \( C^0(X) \):

\[
d\mathcal{F}_{\mu_0} = e^{\beta u} \mu_0
\]

The latter equation is the critical point equation for

\[
G_{\mu_0, \beta}(u) := \mathcal{F}(u) - \frac{1}{\beta} \int e^{\beta u} \mu_0
\]

Theorem 3.6. There exists a unique continuous solution \( u \) to the equation (3.6).

The corresponding probability measure \( \mu := d\mathcal{F}_{\mu_0} \) is the unique minimizer of the corresponding free energy functional \( F_\beta \) on \( \mathcal{M}_1(X) \).

Proof. The existence can be obtained using a variant of the balanced functions used in the proof of Theorem 3.3. One simply replaces the integration measure \( \mu_0 \) in the previous definitions with the measure \( \mu_u := e^{\beta u} \mu_0 \) (followed by a suitable normalization). For example, one sets

\[
\pi_N(u)(x) := \frac{1}{\beta_N} \log \left( 1 - \frac{\beta}{\beta_N} \right) \int_{y \in X^{N-1}} e^{-(\beta_N H_N(x, y) + u(x+y))} (\mu_u^\otimes (N-1)),
\]

where now \( Z[u] := \int_X e^{-(\beta_N H_N(x, y) + u(x))} \mu_u^\otimes N \). Then a direct calculation gives

\[
d\mathcal{F}_{\mu_u}^{(N)} = e^{\beta_N (\pi_N(u) - u)} \mu_u
\]

where now \( \mathcal{F}_{\mu_u}^{(N)} \) has been defined wrt the integration measure \( \mu_u \) and hence \( \pi_N(u) = u \) iff \( d\mathcal{F}_{\mu_u}^{(N)} = \mu_u \) i.e. iff \( u \) is a critical point of \( G_{\mu_u, \beta}(u) \) or equivalently (by concavity) \( u \) is a maximizer of \( G_{\mu_u, \beta}(u) \). Moreover, this time we can directly apply
the Banach fixed point theorem on $C^0(X)$ to obtain a fixed point of $\pi_N$. Indeed, $\pi_N$ defines a contraction mapping on $C^0(X)$ equipped with the sup-norm:

$$||\pi_N(u) - \pi_N(v)|| \leq (1 - \beta/\beta_N) ||u - v||.$$ 

To see this note that, for any constant $c$, $\pi_N(u + c) = \pi_N(u) + (1 - \beta/\beta_N)$ and if $u \leq w$ then $\pi_N(u) \leq \pi_N(w)$. Taking $c := ||u - v||$ and $w := u + c$ thus proves the upper bound in the contraction property above and hence, by symmetry, the lower bound as well (interchanging $u$ and $v$).

Finally, to prove that $\mu_* := dF/du$ minimizes the free energy functional $F_\beta$ on $M_1(X)$ we take $\mu$ in $M_1(X)$ with finite entropy and consider the affine segment $\mu_t := (1 - t)\mu_0 + t\mu$. By basic properties of Legendre transforms (compare Lemma 4.5), if $u_\mu$ is a potential for $\mu$, then $-u_{\mu}$ is a sub-differential for $E$ at $\mu$ and in particular

$$dE(\mu_t) dt = \geq \int (-u)(\mu - \mu_*).$$

Recall also the well-known fact that the differential of the relative entropy $D_{\mu_0}$ (at a point $\mu$ in the convex set where it is finite) is represented by $\log(\mu/\mu_0)$ (using for example that $D_{\mu_0}$ is the Legendre transform of $u \mapsto \log \int e^u \mu_0$). In particular,

$$\frac{dF_\beta(\mu_t)}{dt} |_{t=0} \geq \int \left(-u + \frac{1}{\beta} \log(\mu_*/\mu_0)\right) (\mu - \mu_*) = 0$$

and hence by convexity $F_\beta(\mu) \geq F_\beta(\mu_*)$, as desired. \hfill $\square$

4. THE LARGE DEVIATION PRINCIPLE FOR GIBBS MEASURES

4.1. Setup: the Gibbs measure $\mu^{(N)}_\beta$ associated to the Hamiltonian $H^{(N)}$.

Let $X$ be topological space assumed to be compact (occasionally we will also consider cases where $X$ is non-compact, in particular in the Monge-Ampère setting). A random point process with $N$ particles is by definition a probability measure $\mu^{(N)}$ on the $N$–particle space $X^N$ which is symmetric, i.e. invariant under permutations. Its one point correlation measure $\mu^{(N)}_1$ (or first marginal) is the probability measure on $X$ defined as the push forward of $\mu^{(N)}_1$ to $X$ under the map $X^N \to X$ given by projection onto the first factor (or any factor, by symmetry):

$$\mu^{(N)}_1 := \int_{X^{N-1}} \mu^{(N)}$$

(similarly, the $j$–point correlation measure $\mu^{(N)}_j$ is defined as the $j$ th marginal, i.e. the push forward to $X^j$). In the following we will denote by $M_1(Y)$ the space of all probability measures on a space $Y$ and we will be particularly concerned with the case when $Y = X^N$. In the latter case we will usually use the notation $\mu_N$ for (not necessarily symmetric) elements of $M_1(\mu_N)$ and reserve the notation $\mu^{(N)}$ for specific Gibbs measures defined as below.

4.1.1. The canonical Gibbs ensembles associated to the Hamiltonian $H^{(N)}$.

Fix a background probability measure $\mu_0$ with support $X$ and let $H^{(N)}$ be a given $N$–particle Hamiltonian, i.e. a symmetric continuous and bounded function on $X^N$ satisfying the assumptions in section 2.1. Also fixing a positive number $\beta$ the corresponding Gibbs measure is the symmetric probability measure on $X^N$ defined as

$$\mu^{(N)}_{\beta} := e^{-\beta H^{(N)}} \mu_0^{\otimes N}/Z_N,$$
The normalizing constant
\[ Z_{N,\beta} := \int_{X^N} e^{-\beta H^{(N)}} \mu_0^\otimes N \]
is called the \((N \text{-particle})\) partition function. Occasionally we will simplify the notation by omitting the subscript \(\beta\).

### 4.2. Mean entropy, energy and free energy.

First we recall the general definition of the relative entropy (or the Kullback–Leibler divergence) of two measures \(\nu_1\) and \(\nu_2\) on a space \(Y\) : if \(\nu_1\) is absolutely continuous with respect to \(\nu_2\), i.e. \(\nu_1 = f \nu_2\), one defines
\[ D(\nu_1, \nu_2) := \int_Y \log(\nu_1/\nu_2) \nu_1 \]
and otherwise one declares that \(D(\mu) := \infty\). Note the sign convention used: \(D\) is minus the physical entropy.

Next, we define the mean entropy (relatively \(\mu_0^\otimes N\)) of a probability measure \(\mu_N\) on \(X^N\) (i.e. \(\mu_N \in \mathcal{M}_1(X^N)\)) as
\[ D^{(N)}(\mu_N) := \frac{1}{N} D(\mu_N, \mu_0^\otimes N). \]

When \(N = 1\) we will simply write \(D(\mu) := D^{(1)}(\mu) = D(\mu, \mu_0)\). On the other hand the mean energy of \(\mu_N\) is defined as
\[ E^{(N)}(\mu_N) := \frac{1}{N} \int_{X^N} H^{(N)} \mu_N \]

Finally, the mean (Gibbs) free energy functional on \(\mathcal{M}_1(X^N)\) is now defined as
\[ F^{(N)} := E^{(N)} + \frac{1}{\beta} D^{(N)} \]

Next, we will collect some basic general lemmas. First we have the following simple special case of the well-known sub-additivity of the entropy.

**Lemma 4.1.** The following properties of the entropy hold:
- \(D(\nu_1, \nu_2) \geq 0\) with equality iff \(\nu_1 = \nu_2\)
- For a product measure on \(X^N\)
  \[ D^{(N)}(\mu^\otimes N) = D(\mu) \]

  More generally,
  \[ D^{(N)}(\mu_N) \geq D(\mu_{N,1}), \]
  where \(\mu_{N,1}\) is the corresponding first marginal (one point correlation measure) on \(X\).

The proof of the previous lemma uses only the (strict) concavity of the function \(t \mapsto \log t\) on \(\mathbb{R}\) (see for example [30]). The latter (strict) concavity also immediately gives the following

**Lemma 4.2. (Gibbs variational principle).** Fix \(\beta > 0\). Given a function \(H^{(N)}\) on \(X^N\) and a measure \(\mu_0\) on \(X\), the corresponding free energy functional \(F^{(N)}\) on \(\mathcal{M}_1(X^N)\) attains its minimum value on the corresponding Gibbs measure \(\mu^{(N)}_\beta\) and only there. More precisely,
\[ \inf_{\mathcal{M}_1(X^N)} \beta F^{(N)} = \beta F^{(N)}(\mu^{(N)}_\beta) = -\frac{1}{N} \log Z_N \]
Proof. We recall the simple proof: since $\log(ab) = \log a + \log b$, we have

$$F^{(N)}(\mu_N) = \frac{1}{\beta N} \int_{X^N} \log(\mu_N/e^{-\beta H^{(N)}(\mu \otimes \mu)N}) \mu_N := \frac{1}{\beta N} \int_{X^N} \log(\mu_N/\mu^{(N)}) \mu_N - \frac{1}{N\beta} \log Z_N$$

which proves the lemma using Jensen’s inequality (i.e. the first point in the previous lemma).

Note that the same argument applies if $\beta < 0$, since we can simply replace $H^{(N)}$ with $-H^{(N)}$ in the previous argument (as long as the corresponding partition function $Z_N$ is finite).

4.3. Convergence of the one-point correlation measure towards the minimizer of the free energy. Given a background measure $\mu_0$ on $X$ and a Hamiltonian $H^{(N)}$ satisfying the assumptions in section 2.1 we define, for any $\beta > 0$, the corresponding free energy functional $F_{\beta}$ on $M_1(X)$:

$$F_{\beta}(\mu) := E(\mu) + D_{\mu_0}(\mu)/\beta,$$

where $E(\mu)$ is the thermodynamical energy defined by 2.5. We start by proving a weak version of Theorem 1.4, stated in the introduction.

**Theorem 4.3.** Let $H^{(N)}$ be an $N$-particle Hamiltonian on $X^N$ satisfying the assumptions in section 2.1 we define, for any $\beta > 0$, the corresponding free energy functional $F_{\beta}$ on $M_1(X)$:}

$$\lim_{N \to \infty} -\frac{1}{\beta N} \log Z_{N,\beta} = \inf_{\mu \in M_1(X)} F_{\beta}$$

Take a measure $\mu \in M_1(X)$. By the Gibbs variational principle (i.e. the previous lemma) we have

$$F^{(N)}(\mu^{(N)}) \leq F^{(N)}(\mu^{(N)}) = \frac{1}{N} \int H^{(N)}(H^{(N)} + u) \mu^{(N)} - \frac{1}{\beta} D(\mu),$$

where we have used the first point in Lemma 4.1 in the last equality. Hence, applying the upper bound in Theorem 2.1 gives

$$F^{(N)}(\mu^{(N)}) \leq \inf_{\mu \in M_1(X)} (E(\mu) + D(\mu))$$

To obtain a lower bound we apply the second point in Lemma 4.1 to get

$$\frac{1}{N} \int H^{(N)}(H^{(N)} + u) \mu^{(N)} - \frac{1}{\beta} D(\mu) \leq F^{(N)}(\mu^{(N)})$$

Next, we fix $u \in C^0(X)$ and rewrite the first term in the lhs above as follows:

$$\frac{1}{N} \int H^{(N)}(H^{(N)} + u) \mu^{(N)} = \int \frac{1}{N} (H^{(N)} + u) \mu^{(N)} - \int_X u \mu_1^{(N)}$$

and hence replacing the first integral in the rhs with its infimum gives

$$\inf_{\chi^{N}} \frac{1}{N} (H^{(N)} + u) - \int_X u \mu_1^{(N)} \leq \frac{1}{N} \int H^{(N)} \mu^{(N)},$$
which by definition means that, if \( \mu_* \) denotes a weak limit point of the sequence \( \mu_1^{(N)} \), then
\[
F(u) - \int_X u \mu_* \leq \liminf_{N \to \infty} \frac{1}{N} \int H^{(N)} \mu^{(N)}.
\]
Taking the sup over all \( u \in C^0(X) \) thus gives
\[
E(\mu_*) \leq \liminf_{N \to \infty} \frac{1}{N} \int H^{(N)} \mu^{(N)}.
\]
Next, since \( D \) is lower-semi-continuous we have
\[
D(\mu_*) \leq \liminf_{N \to \infty} D(\mu_1^{(N)}).
\]
All in all, this means that
\[
F(\mu_*) \leq \liminf_{N \to \infty} F^{(N)}(\mu \otimes N) \leq \limsup_{N \to \infty} F^{(N)}(\mu \otimes N) \leq \inf_{\mu \in \mathcal{M}_1(X)} F_\beta(\mu)
\]
But then it must be that all the inequalities in the previous line are actually equalities. In particular, \( \mu_* \) is a minimizer of \( F \).

4.4. **Proof of the Large Deviation Principle (Theorem 1.4).** Let us recall the general definition of a LDP due to Donsker and Varadhan (see for example the book [24]):

**Definition 4.4.** Let \( \mathcal{M} \) be a Polish space, i.e. a complete separable metric space.

(i) A function \( I : \mathcal{M} \to [0, \infty] \) is a rate function iff it is lower semi-continuous. It is a good rate function if it is also proper.

(ii) A sequence \( \Gamma_k \) of measures on \( \mathcal{M} \) satisfies a large deviation principle with speed \( r_k \) and rate function \( I \) if
\[
\limsup_{k \to \infty} \frac{1}{r_k} \log \Gamma_k(\mathcal{F}) \leq -\inf_{\mu \in \mathcal{F}} I(\mu)
\]
for any closed subset \( \mathcal{F} \) of \( \mathcal{M} \) and
\[
\liminf_{k \to \infty} \frac{1}{r_k} \log \Gamma_k(\mathcal{G}) \geq -\inf_{\mu \in \mathcal{G}} I(\mu)
\]
for any open subset \( \mathcal{G} \) of \( \mathcal{M} \).

4.4.1. **Functional analytic framework and Legendre-Fenchel transforms.** Let \( X \) be a Polish space and denote by \( C_b(X) \) the space of all bounded continuous functions on \( X \) and by \( \mathcal{M}(X) \) the space of all signed finite Borel \( \mu \) measures on \( X \). We will write the corresponding integration pairing as
\[
\langle u, \mu \rangle := \int_X u \mu
\]
We equip \( \mathcal{M}(X) \) with the weak topology generated by \( C_b(X) \). Then we may identify \( C_b(X) \) with the topological dual \( \mathcal{M}(X)^* \) of \( \mathcal{M}(X) \), i.e. with the space of all linear continuous functions on \( \mathcal{M}(X) \) [24]. We will be mainly concerned with the subspace \( \mathcal{M}_1(X) \) of all probability measures on \( X \) which is a convex subset of \( \mathcal{M}(X) \) (and compact iff \( X \) is compact). This latter space is a locally convex topological vector
space. As such it admits a good duality theory (see section 4.5.2 in [24]): given a functional $\Lambda$ on the vector space $C_b(X)$ its Legendre(-Fenchel) transform is the following functional $\Lambda^*$ on $\mathcal{M}(X)$:

$$
\Lambda^*(\mu) := \sup_{u \in C_b(X)} (\Lambda(u) - \langle u, \mu \rangle)
$$

Conversely, if $H$ is a functional on the vector space $\mathcal{M}(X)$ we let

$$
H^*(u) := \inf_{\mu \in \mathcal{M}(X)} (H(\mu) + \langle u, \mu \rangle)
$$

Note that we are using rather non-standard sign conventions. In particular, $\Lambda^*(\mu)$ is always convex and lower semi-continuous (lsc), while $H^*(u)$ is concave and upper-semicontinuous (usc). As a well-known consequence of the Hahn-Banach separation theorem we have the following fundamental duality relation (Lemma 4.5.8 in [24]):

\begin{equation}
\Lambda = (\Lambda^*)^*
\end{equation}

iff $\Lambda$ is concave and usc. We also recall the following standard

**Lemma 4.5.** Assume that $\Lambda$ is a functional on $C_b(X)$ which is finite, lsc, concave and Gateaux differentiable (i.e differentiable along lines). Then, for a fixed $u \in C_b(X)$ the differential $d\Lambda|_u$ is the unique minimizer of the following functional on $\mathcal{M}(X)$:

\begin{equation}
\mu \mapsto \Lambda^*(\mu) + \langle u, \mu \rangle
\end{equation}

(and the minimum value equals $\Lambda(u)$). Conversely, if the latter functional admits a unique minimizer $\mu_u$ for any $u \in C_b(X)$, then $\Lambda(u)$ is Gateaux differentiable and $d\Lambda|_u = \mu_u$.

Recall that, in general, if $\Gamma$ is a probability measure on a topological vector space $\mathcal{M}$ then its Laplace transform is the following functional defined on the topological dual $\mathcal{M}^*$ of $\mathcal{M}$:

$$
\hat{\Gamma}[u] := \int_{\mu \in \mathcal{M}} \Gamma e^{-(u, \mu)}
$$

(assuming that the integral is finite).

We will use the following abstract form of the Gärtner-Ellis theorem (see [24], Cor 4.6.14, p. 148) and references therein about the different versions of this theorem) which can be seen as an infinite dimensional version of the method of stationary phase, i.e. the Laplace method.

**Theorem 4.6.** (abstract Gärtner-Ellis Theorem). Let $\mathcal{M}$ be a locally convex Hausdorff topological vector space and $\Gamma_k$ a sequence of Borel measures on $\mathcal{M}$ which is exponentially tight with respect a sequence $r_k$ of positive numbers and such that the Laplace transforms $\hat{\Gamma}_k$, seen as functionals on the dual $\mathcal{M}^*$,

\begin{equation}
-\frac{1}{r_k} \log \hat{\Gamma}_k[r_k u] \to \Lambda[u]
\end{equation}

for any $u$ in $\mathcal{M}^*$ where the functional $\Lambda$ is Gateau differentiable on $\mathcal{M}^*$. Then $\Gamma_k$ satisfies a LDP with speed $r_k$ and with a rate functional $H := \Lambda^*$ on $\mathcal{M}$, i.e. the rate function $H$ is the Legendre-Fenchel transform of $\Lambda$.

The definition of exponential tightness will be recalled in the proof of Theorem 1.1 (in case $X$ is compact this condition is automatically satisfied).
4.4.2. End of proof of Theorem 1.4. Given \( u \in C^0(X) \) we let \( f_N[u] := -\frac{1}{N\beta} \log \int \mu(N)e^{-N\beta(u, \delta_N)} \), i.e. the scaled logarithm of the Laplace transform at \( u \) of the probability measure \( \Gamma := (\delta_N)_* \mu(N) \) on \( \mathcal{M}(X) \). It may also be written as

\[
\Lambda_N[u] = -\frac{1}{N\beta} \log \frac{Z_{N, \beta}[u]}{Z_{N, \beta}[0]},
\]

where

(4.4) \[
Z_{N, \beta}[u] = \int e^{-\beta(H + u)} \mu_0^N
\]

By the previous theorem applied to \( H(N)u := H(N) + u \) we have

\[
\Lambda_N[u] \to \Lambda(u) := \inf_{\mu \in \mathcal{M}_1(X)} (E(\mu) + \int_X u \mu + \frac{1}{\beta} D(\mu) - C) := \inf_{\mu \in \mathcal{M}_1(X)} (F_\beta(\mu) + \int_X u \mu - C);
\]

Next, we observe that the infimum in the rhs above is up to a harmless additive constant, by definition, the Legendre transform of the extended free energy functional \( \mu \mapsto F_\beta(\mu) \), defined by the expression (4.1) on \( \mathcal{M}_1(X) \) and set to be equal to \( \infty \) on the complement of \( \mathcal{M}_1(X) \) in \( \mathcal{M}(X) \). But, as explained in the proof of the previous theorem, the corresponding functional is strictly convex on the subspace where it is finite. But then it follows from Lemma 4.5 (and the duality relation \( F = \Lambda^* \)) that \( \Lambda \) is Gateaux differentiable and hence we can apply the Gärtner-Ellis theorem to conclude.

4.5. The ambient point of view. To better see the connection to the Monge-Ampère setting, to be considered in the following section, we consider the following general “ambient setting”. Start with a (possible non-compact) Riemannian manifold \( (Y, g) \) - to be referred to as the “ambient space” (which in the Monge-Ampère will be equal to \( \mathbb{R}^n \) equipped with the Euclidean metric) and a Hamiltonian \( H(N) \) on \( Y \) satisfying the Lipschitz assumption in section 2.1, with a Lipschitz constant \( L \), defined with respect to the Riemannian metric \( g \). Fix a reference element \( \phi_0 \) in the corresponding space \( \mathcal{P}(Y) \) (to be referred to as the space of “ambient potentials”) and assume that

- For any function \( u \in C_b(Y) \) the following limit exists

\[
\mathcal{F}(u) := \lim_{N \to \infty} \inf_{\mu \in \mathcal{M}_1(X)} \frac{1}{N}(H(N) + \phi_0 + u),
\]

- The functional \( \mathcal{F}(u) \) is Gateaux differentiable.

Also, on the subspace \( \mathcal{P}_+(Y) \) of all functions \( \phi \in \mathcal{P}(Y) \) such \( \phi - \phi_0 \) is bounded we can then define an operator

\[
M : \mathcal{P}(Y) \to \mathcal{M}_1(Y), \quad M(\phi) = d\mathcal{F}_{\phi - \phi_0}
\]

(which can be seen as a generalized Monge-Ampère operator).

Now, given a sequence \( \beta_N \) of positive numbers any choice of probability measure \( \mu_0 \) on \( Y \) with compact support \( X \) induces a random point processes on \( X \) defined by the Gibbs measure of the restricted Hamiltonian (compare section 4.1). Note that, by construction, there is a natural extension map from \( \mathcal{P}_N(X) \) to \( \mathcal{P}_N(Y) \). Under suitable assumptions one can then show that the sequence of balanced metrics appearing in the proof of Theorem 3.6 converge to a function \( \phi \) satisfying the following generalized mean field equations on the ambient space \( Y \):
We will not develop this general theory further, as the corresponding results will be obtained directly in the particular Monge-Ampère setting considered in section 5.5. A remarkable feature of the latter setting is that the corresponding operator $M(\phi)$ is a local operator, since it coincides with the Monge-Ampère measure which (up to regularization of metric functions on $X$) is a differential operator.

4.6. An alternative proof of the LDP for equicontinuous Hamiltonians on compact spaces. In this section we assume again that $X$ is a compact metric space equipped with a probability measure $\mu_0$ whose support is equal to $X$. Let us start by recalling a large deviation result from [25], which in the present setting may be formulated as follows.

**Theorem 4.7.** [25] Let $X$ be a compact metric space, $H_N$ a sequence of bounded continuous symmetric functions on $X^N$ and $U$ a bounded continuous function on $\mathcal{M}_1(X)$ such that $\sup_{X^N} |H^{(N)}/N - U(\delta_N)| \to 0$, as $N \to \infty$. Then, for any real number $\beta$ the measure $(\delta_N)_*(e^{-\beta H^{(N)}/\mu_0^{\otimes N}})$ on $\mathcal{M}_1(X)$ satisfies a LDP with speed $N$ and rate functional $\beta E + D_{\mu_0}$.

As will be explained below the proof can be reduced to the special case when $H^{(N)} = 0$, which is the content of Sanov’s classical result (this is slightly different than the reduction to Sanov’s theorem in [25], which uses a Laplace principle). Using Theorem 4.7 we will in this section give a more direct proof of Theorem 1.4 stated in the introduction. In fact, we will obtain the following more general form of the latter theorem 1.4.

**Theorem 4.8.** Let $X$ be a compact metric space and $H^{(N)}/N$ a sequence of symmetric functions on $X^N$ which is uniformly bounded and equicontinuous (in the sense of section 2.1). Let $\beta_N$ be a sequence of positive number tending to infinity and assume that $(\delta_N)_*(e^{-\beta_N H^{(N)}/\mu_0^{\otimes N}})$ satisfies an LDP with rate functional $E(\mu)$ and speed $\beta_N N$. Then, for $\beta$ any fixed (possibly negative) number $(\delta_N)_*(e^{-\beta H^{(N)}/\mu_0^{\otimes N}})$ satisfies a LDP with speed $\beta N$ and rate functional $E + D_{\mu_0}/\beta$.

The previous result combined with the Gärtner-Ellis theorem immediately gives Theorem 1.4 stated in the introduction. Note that in the setting of the latter theorem the functional $E$ is automatically convex. In order to reduce the proof of the previous theorem to Theorem 4.7 we will invoke the following

**Lemma 4.9.** There exists a continuous and bounded function $U$ on $\mathcal{M}_1(X)$ such that $\lim_{N \to \infty} \sup_{X^N} |H_{N_j}/N_j - (\delta_{N_j})^* U| = 0$. In fact, $U$ can be taken to have the same modulus of continuity as the sequence $H^{(N)}/N$.

**Proof.** First we may, using the map $\delta_N$ embed $X^{(N)}$ isometrically in $\mathcal{M}_1(X)$, equipped with the Wasserstein 1-metric. Then $H^{(N)}/N$ extends to a continuous function $U_N$ defined on all of $\mathcal{M}_1(X)$ preserving the modulus of continuity. Accepting this for the moment the existence of $U$ follows immediately from the Arzelà-Ascoli theorem applied to the sequence $U_N$ on the compact space $\mathcal{M}_1(X)$. As for the extension property if follows from general considerations: let $\mathcal{K}$ be a subspace
of a metric space $(\mathcal{M}, d)$ and $u_N$ a sequence of functions on $\mathcal{K}$ which we for simplicity assume is uniformly Lipschitz continuous (the general case is proved in a similar manner). Setting $U_N(y) := \inf_{x \in \mathcal{K}} (u(x) + d(x, y))$ then gives the desired extension. 

To prove Theorem 4.8 we start by showing that the following claim: $E(\mu) = U(\mu)$

holds, for any $\mu$ such that $D_{\mu_0}(\mu) \neq 0$ and in particular the convergence in the previous lemma must hold for the full sequence indexed by $N$. To prove the claim first note that the assumed LDP implies, by a general result for LDPs (see Theorem 4.1.18 in [24]), that, setting $E$\(=\) \(e^{-\beta_N H^N(\mu_0^N)}\),

\[
\lim_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{\beta_N N} \log \int_{B_\delta(\mu)} \Gamma_N,\beta_N = \lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{\beta_N N} \log \int_{B_\delta(\mu)} \Gamma_N,\beta_N = E(\mu),
\]

where $B_\delta(\mu)$ is the ball of radius $\delta$ centered at $\mu$, defined with respect to any metric compatible with the weak topology on the space $\mathcal{M}_1(X)$ and since $X$ is assumed compact we can take the metric to the Wasserstein 1-metric. Now, by Sanov’s theorem

\[
\lim_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \int_{B_\delta(\mu)} (\delta_N)_* \mu_0^N \geq -D_{\mu_0}(\mu)
\]

Hence, if $D_{\mu_0}(\mu) < \infty$ the previous lemma gives (again using that the map \(2.3\) defined by $\delta_N$ is an isometry) that

\[U(\mu) + \frac{1}{\beta_N} O(1) = E(\mu),\]

where $O(1)$ denotes a bounded term and since $\beta_N \to \infty$ this proves the claim above. We can now apply Theorem 4.7 to deduce the LDP with respect to a fixed $\beta$ with the rate functional $\beta U + D$ and speed $N$, which according to the previous claim coincides with $\beta E + D$, as desired (using that $D_{\mu_0}(\mu) = \infty$ iff $F_\beta(\mu) = \infty$). Alternatively, repeating the arguments above with $\beta_N$ replaced with $\beta$ immediately gives, using the previous lemma, that

\[
\lim_{\delta \to 0} \liminf_{N \to \infty} \frac{1}{\beta N} \log N \int_{B_\delta(\mu)} \Gamma_N,\beta = \lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{\beta N} \log N \int_{B_\delta(\mu)} \mu_0^N = U + \frac{1}{\beta} D,
\]

which coincides with $E + \frac{1}{\beta} D$. Finally, by Theorem 4.1.11 and Lemma 1.2.18 in [24] the desired LDP holds follows.

5. Permanental point processes, the Monge-Ampère and optimal transport

5.1. Setup. Consider $\mathbb{R}^n$ equipped with the Euclidean scalar product that we will denote by a dot. The corresponding Euclidean coordinates will be denoted $x = (x_1, ..., x_n)$. We will identify the dual linear space with $\mathbb{R}^n$ equipped with the coordinates $p = (p_1, ..., p_N)$ so that the corresponding duality pairing may be written as $\langle x, p \rangle = x \cdot p$. We fix once and for all a convex body $P$ in $\mathbb{R}^n$. Without loss of generality we may assume that 0 is contained in the interior of $P$ and that $P$ has unit Euclidean volume. The support function of $P$ will be denoted by $\phi_P$, i.e. $\phi_P(x) := \sup_{p \in P} \langle x, p \rangle$. The set of all bounded continuous function $u$ on $\mathbb{R}^n$ will be denoted by $C_b(\mathbb{R}^n)$. 

25
Given a convex function $\phi$ on $\mathbb{R}^n$, we will denote by $MA(\phi)$ its Monge-Ampère measure (in the sense of Alexandrov), i.e., if $E$ is a Borel set then $(MA(\phi))(E)$ is defined as the Lebesgue measure of the image of $E$ under the sub-gradient $\nabla \phi$ of $\phi$ (viewed as a multivalued map from $\mathbb{R}^n$ to $\mathbb{R}^n$) [27]. Following [9] we will denote by $\mathcal{P}(\mathbb{R}^n)$ the space of all convex functions $\phi$ on $\mathbb{R}^n$ such that $\phi - \phi_P$ is bounded from above and by $\mathcal{P}_+(\mathbb{R}^n)$ the subspace of all $\phi$ such that $\phi - \phi_P$ is bounded.

Given a (possible non-convex) upper-semi continuous (usc) function $\phi$ we will write $\phi^*$ for its Legendre transform, i.e.

$$\phi^*(p) := \sup_{x \in \mathbb{R}^n} \langle x, p \rangle - \phi(x)$$

By basic properties of the Legendre transform we have that, if $\phi$ is in $\mathcal{P}(\mathbb{R}^n)$, then the subgradient image $(\nabla \phi)(\mathbb{R}^n)$ is contained in $P$ and $\phi^*$ is equal to $\infty$ on the complement of $P$. Similarly, under the Legendre transform the space $\mathcal{P}_+(\mathbb{R}^n)$ corresponds to the space of all bounded convex functions on $P$ (see [9] and references therein). Finally, following [9] we will say that $\phi$ $\mathcal{P}(\mathbb{R}^n)$ has finite energy if $\phi^*$ is integrable on $P$, i.e., if

$$\mathcal{E}(\phi) := - \int_P \phi^*(p)dp > -\infty$$

and we will denote the space of all finite energy convex functions by $\mathcal{E}_P^1(\mathbb{R}^n)$. We thus have the following (strict) inclusions:

$$\mathcal{P}_+(\mathbb{R}^n) \subset \mathcal{E}_P^1(\mathbb{R}^n) \subset \mathcal{P}(\mathbb{R}^n)$$

Let us finally recall the following basic compactness property of the space $\mathcal{P}(\mathbb{R}^n)$ which is an immediate consequence of the Arzelà-Ascoli theorem (compare [27]).

**Proposition 5.1.** Let $\phi_j$ be a sequence of normalized functions in $\mathcal{P}(\mathbb{R}^n)$, which by definition means that $\sup_{\mathbb{R}^n} (\phi_j - \phi_P) = 0$ or equivalently that $\phi_j(0) = 0$. Then, perhaps after passing to a subsequence, $\phi_j$ converges locally uniformly to a normalized function $\phi$ in $\mathcal{P}(\mathbb{R}^n)$.

5.1.1. **Weighted sets and measures.** By definition a weighted set $(X, \phi_0)$ consists of a closed set $X$ in $\mathbb{R}^n$ and a weight function $\phi_0$ on $X$, i.e., a continuous function $\phi_0$ on $X$ such that $\phi - \phi_P \to \infty$ as $|x| \to \infty$ in $X$ (in particular if $X$ is compact then latter growth condition is vacuous). Occasionally we will identify $\phi_0$ with a function on $\mathbb{R}^n$ by letting the extension be identically equal to $\infty$ on the complement of $X$.

Given a measure $\mu_0$ on $X$ we say that $(\mu_0, \phi_0)$ is a weighted measure if $\phi_0$ is a weight function on the support $X$ of $\mu_0$ and $e^{\beta(\phi_P - \phi_0)} \mu_0$ has finite total mass for any positive number $\beta$.

Given a closed set $X$ we define the corresponding projection operator $\Pi_X$ from the space of weights on $X$ to the space $\mathcal{P}_+(\mathbb{R}^n)$ by the following convex envelope

$$\Pi_X(\phi_0)(x) = \sup_{\phi \in \mathcal{P}(\mathbb{R}^n)} \{ \phi(x) : \phi \leq \phi_0 \text{ on } X \}$$

(if $\phi$ is continuous then $\Pi_X(\phi)$ is in $\mathcal{P}(\mathbb{R}^n)$ and if $\phi$ is a weight function then $\Pi_X(\phi)$ is in $\mathcal{P}_+(\mathbb{R}^n)$, since $\phi_P - C$ is a candidate for the sup above, if $C$ is sufficiently large). In particular, if $(X, \phi_0)$ is a weighted set we will occasionally write $\phi_e := \Pi_X(\phi_0)$ and $\mu_e := MA(\Pi_X(\phi_0))$. From the growth assumption on $\phi_0$ it follows that the incidence set

$$D_0 := \{ \Pi_X(\phi_0) = \phi_0 \}$$
is compact and, since, by general properties of free convex envelopes, $\text{MA}(\Pi_X(\phi))$ is always contained in $D_0$ we conclude that $\mu_\star$ has compact support contained in $X$. We also note that “the Legendre transform doesn’t see the projection $\Pi_X$” in the following sense

\begin{equation}
(\Pi_X(\phi_0))^\star = \phi_0^\star \text{ on } P,
\end{equation}

(which is a special case of formula \[5.3\] in Lemma \[5.2\] below). We note that $\Pi_X(\phi_0)$ is bounded since $\Pi_X(\phi_0)$ is in $\mathcal{P}_+(\mathbb{R}^n)$.

5.1.2. The generalized permanental point processes. Set

$$\text{Per}(x_1, \ldots, x_{N_\beta}) := \text{Per}(e^{x_1 \cdot p_j}),$$

where $p_j$ ranges over the $N$ lattice points in $kP$ (compare the notation in the introduction of the paper). For a given sequence $\beta_N$ such that $\lim_{N \to \infty} \beta_N = \beta \in [0, \infty]$ we let $\mu_{\beta_N}^{(N)}$ be the probability measure on $X^N$ defined by

$$\mu_{\beta_N}^{(N)} := \frac{1}{Z_{N,\beta_N}} \text{Per}(x_1, \ldots, x_{N_\beta})^{\beta_N / k} e^{-\beta_N (\phi(x_1) + \cdots + \phi(x_N))} \mu_0^\otimes N,$$

where $Z_{N,\beta_N}$ is the normalizing constant (which is finite by the growth assumption on $\phi_0$; see below). Setting

\begin{equation}
H_{\beta_0}^{(N)} := -\frac{1}{k} \text{Per}(x_1, \ldots, x_{N_\beta}) + \phi(x_1) + \cdots + \phi(x_N)
\end{equation}

the probability measure $\mu_{\beta_N}^{(N)}$ thus becomes the Gibbs measure, at inverse temperature $\beta_N$, determined by the $N$-particle Hamiltonian $H_{\beta_0}^{(N)}$.

5.2. Large $N$ asymptotics and variational principles. Since the logarithm of any convex combination of functions of the form $e^{x \cdot p}$ for $p \in P$ is in the space $\mathcal{P}(\mathbb{R}^n)$ it follows that for $(x_2, \ldots, x_N) \in X^N$ fixed $\psi(x) := -H_{\phi_0}^{(N)}(x, x_2, \ldots, x_N)$ defines an element in $\mathcal{P}(\mathbb{R}^n)$ and similarly when the other variables are fixed. In particular, the partial gradients satisfy

\begin{equation}
\nabla_{x_i} H_{\phi_0}^{(N)} \in -P
\end{equation}

for any $N$ and any $x_i \in X$. Since $P$ is assumed bounded $H_{\phi_0}^{(N)}/N$ is uniformly Lipschitz continuous \[2.4\] and when $X$ is compact $H_{\phi_0}^{(N)}/N$ is thus equicontinuous in the sense of section \[2.1\]. In order to be able to handle the non-compact case we recall that the support of $\mu_\star := \text{MA}(\Pi_X \phi_0)$ has compact support and it is thus contained in a large ball $B_R$.

**Lemma 5.2.** Let $(\mu_0, \phi_0)$ be a weighted measure and $X$ the support of $\mu_0$. Then

$$\Pi_{X \cap B_R} \phi_0 = \Pi_X \phi_0$$

and for any $\phi_k$ in $\mathcal{P}(\mathbb{R}^n)$ we have

\begin{equation}
\sup_X e^{k(\phi_k - \phi_0)} = \sup_{\mathbb{R}^n} e^{k(\phi_k - \Pi_X \phi_0)} = \sup_{D_{\phi_0}} e^{k(\phi_k - \phi_0)}
\end{equation}

Moreover, for any $\epsilon > 0$, there is $C_\epsilon > 0$ (independent of $\phi_k$) such that

$$\sup_X e^{k(\phi_k - \phi_0)} \leq C_\epsilon e^{k(\Pi_X \phi_0 - \phi_0)} e^{k} \int_{X \cap B_R} e^{k(\phi_k - \phi_0)} \mu_0$$
Proof. By definition $\Pi_{X \setminus B_R} \varphi \geq \Pi_X \varphi$ and $\Pi_{X \cap B_R} \varphi \leq \varphi$ on the support of $MA(\Pi_X \varphi)$. Since the latter set is contained in the incidence set $D$ (see above) this means that $\Pi_{X \cap B_R} \varphi \leq \Pi_X \varphi$ a.e. w.r.t $MA(\Pi_X \varphi)$ and hence the inequality holds everywhere according to the domination principle for $MA$ (see the appendix). This shows that $\Pi_{X \cap B_R} \varphi = \Pi_X \varphi$. Next, note that the first equality in (5.5) follows directly from the extremal definition of $\Pi_X$ and the second one follows from the domination principle, as in the previous equality. To prove the inequality in the lemma we first observe that, when $X$ is compact we have

$$\sup_X e^{k(\phi_k - \phi_0)} \leq C e^{k} \int_X e^{k(\phi_k - \phi_0)} \mu_0,$$

using that $\phi_k$ is uniformly continuous on $X$ with a constant of continuity only depending on $P$ (in fact, since $\nabla \phi_k$ is in $P$ we even have a Lipschitz constant only depending on $P$). Indeed, given $\epsilon > 0$ and $\phi_k$ we simply estimate $\int_X e^{k(\phi_k - \phi_0)} \mu_0$ from below by the integral over a small poly-disc $\Delta_\delta(x_k)$ of radius $\delta$ centered at the point $x_k \in X$ where the sup of $\phi_k$ is attained. By the Lipschitz property $\delta$ can be chosen so that $e^{k(\phi_k - \phi_0)} \geq e^{k(\phi_k - \phi_0)(x_k) - \epsilon}$ and hence the desired estimate holds with the constant $C_\delta := \inf_{x \in X} \mu_0(\Delta_\delta(x))$, which is strictly positive for any $\delta$. Indeed, by the definition the support $X$ of a measure $\mu_0$ the continuous function $\mu_0(\Delta_\delta(x))$ on $X$ is point-wise strictly positive, hence globally strictly positive on $X$, by compactness).

In the general case when $X$ may be non-compact we can apply the previous inequality to $X \cap B_R$ and set $\psi := \phi_k - \frac{1}{k} \log(C e^k \int_{X \cap B_R} e^{k(\phi_k - \phi_0)} \mu_0)$ so that $\psi \leq \phi_0$ on $X \cap B_R$. Then, by the extremal definition of $\Pi_{X \cap B_R} \phi_0$, we get $\psi \leq \Pi_{X \cap B_R} \phi_0 = \Pi_X \phi_0$ on all of $\mathbb{R}^n$ which concludes the proof of the lemma. \(\square\)

Proposition 5.3. Let $\mu_0$ be a measure on $\mathbb{R}^n$ with support $X$ and $\phi_0$ a weight function on $\mathbb{R}^n$ such that $e^{\beta(\phi_0 - \phi_0)} \mu_0$ has finite total mass for some non-negative number $\beta$. Then

$$\lim_{N \to \infty} \frac{1}{kN} \log \int_{X^N} P_{\text{Per}}(x_1, \ldots, x_N) e^{-k\phi_0((x_1) + \cdots + \phi_0((x_1)))} \mu_0^{\otimes N} = \int_{\mathbb{R}^n} (\Pi_X \phi_0)^* dp = \int_{\mathbb{R}^n} \phi_0^* dp$$

and the same limit holds when the integral over $X^N$ is replaced with a sup. Equivalently, for any $u \in C_0(\mathbb{R}^n)$ we have

$$-\lim_{N \to \infty} \frac{1}{kN} \log \int_{X^N} e^{-k(H_{\phi_0}^{(N)} + u)} \mu_0^{\otimes N} = \lim_{N \to \infty} \frac{1}{N} \inf_{X^N} (H_{\phi_0}^{(N)} + u) = -\int_{\mathbb{R}^n} (\Pi_X \phi_0)^* dp$$

Proof. First note that

$$\frac{1}{kN} \log \int_{X^N} P_{\text{Per}}(x_1, \ldots, x_N) e^{-k\phi_0((x_1) + \cdots + \phi_0((x_1)))} \mu_0^{\otimes N} = \frac{1}{N} \sum_{p \in \mathcal{P}_{2/k}} \frac{1}{k} \log \int_{\mathbb{R}^n} e^{k(p \cdot x - \phi_0(x))} \mu_0(x) + \log \frac{N!}{N^k}$$

Next, we will prove the lower bound in the proposition. First by the inequality in the previous lemma we may as well, when letting $k \to \infty$, replace the term inside the sum with $v(p) := \sup_{X \cap B_R} (p \cdot x - \phi_0(x))$ to get

$$\frac{1}{kN} \log \int_{X^N} P_{\text{Per}}(x_1, \ldots, x_N) e^{-k\phi_0((x_1) + \cdots + \phi_0((x_1)))} \mu_0^{\otimes N} \geq \frac{1}{N} \sum_{p \in \mathcal{P}_{2/k}} v(p) + o(1)$$

(also using Stirling’s formula, which gives that $\frac{\log N!}{N^k} \to 0$). Moreover, by the previous we may as well replace the sup over $X \cap B_R$ in the definition of $v(p)$ with a
sup over $X$ to get $v = (\Pi \phi_0)^*$ which is bounded on $P$. Hence, since,

\[
(5.6) \quad \frac{1}{N} \sum_{p \in P_{n/k}} \delta_p \to 1_p dp
\]

weakly on $P$, as $k \to \infty$, this concludes the proof of the lower bound:

\[
\lim_{k \to \infty} \frac{1}{kN} \log \int_{X^N} \text{Per}_k(x_1, ..., x_{N_k}) e^{-k\phi_0((x_1) + ... + \phi_0((x_1)))} \mu_0^N \geq \int_P (\Pi X \phi_0)^* dp = \int_P (\phi_0)^* dp
\]

To handle the upper bound let us first, to fix ideas, assume that $\mu_0$ has finite total mass $M$. Then we can trivially estimate

\[
\int_{\mathbb{R}^n} e^{k(p - \phi_0(x))} \mu_0(x) \leq M \sup_{x \in X} e^{k(p - \phi_0(x))}
\]

and conclude as before. To handle the general case we split the integral over $\mathbb{R}^n$ according to the decomposition $\mathbb{R}^n = B_R \cup B_R^\infty$ to get

\[
\int_{\mathbb{R}^n} e^{k(p - \phi_0(x))} \mu_0(x) \leq \sup_{X \cap B_R} e^{k(p - \phi_0(x))} \int_{B_R^\infty} e^{k(P_{X} \phi_0 - \phi_0)} \mu_0(x) \cdot C e^{k} \int_{X \cap B_R} e^{k(\phi_0 - \phi_0)} \mu_0,
\]

using the inequality in the previous lemma in the estimate of the second term, Now, $P_{X} \phi_0 - \phi_0 \leq \phi_P - \phi_0 + C$ and hence, since $\phi_0 - \phi_P \to \infty$ at $\infty$ in $\mathbb{R}^n$ we get $P_{X} \phi_0 - \phi_0 \leq (1 - \delta)(\phi_P - \phi_0)$ for $\delta$ a sufficiently small number (taking $R$ above sufficiently large but fixed). Accordingly, assuming that $e^{\delta x(\phi_P - \phi_0)} \mu_0$ has finite mass $M'$ for some positive number $\beta_0$ we get for $k$ sufficiently large that

\[
\int_{\mathbb{R}^n} e^{k(p - \phi_0(x))} \mu_0(x) \leq \sup_{X \cap B_R} e^{k(p - \phi_0(x))}(M + MC e^{k} \int_{X \cap B_R} \mu_0) \leq C e^{k} \sup_{X \cap B_R} e^{k(p - \phi_0(x))}
\]

Hence, we get, just as before, that

\[
\lim_{k \to \infty} \frac{1}{kN} \log \int_{X^N} \text{Per}_k(x_1, ..., x_{N_k}) e^{-k\phi_0((x_1) + ... + \phi_0((x_1)))} \mu_0^N \leq \int_P (\Pi X \phi_0)^* ,
\]

which concludes the proof of the asymptotics for the integrals in the theorem. Finally, applying the previous lemma to $\phi_k(x) = -\log H(N)(x, x_2, ..., x_N)$ for any choice of $(x_2, ..., x_N) \in X^{N-1}$, etc, one coordinate a time, and arguing as above, also shows that the integral over $X^N$ may as well, asymptotically, be replaced with a sup over $X^N \cap B_R^N$, which in turn coincides with the sup over $X^N$ itself.

5.3. Functional on convex functions and probability measures. Fixing a weighted set $(X, \phi_0)$ we now define, following the notation in section 2.1 a functional $\mathcal{F}(u)$ on $\mathcal{C}_b(\mathbb{R}^n)$ as the limiting functional appearing in Prop 5.3

\[
(5.7) \quad \mathcal{F}(u) := -\int_P (\phi_0 + u)^*(p) dp
\]

The connection to the Monge-Ampère operator appears as follows (compare [9]). Defining

\[
\mathcal{E}(\phi) := -\int_P \phi^*(p) dp
\]

we have $\mathcal{E}(\phi_P) = 0$ and

\[
d\mathcal{E}(\phi) = MA(\phi)
\]
for any \( \phi \in P(\mathbb{R}^n) \) such that \( \mathcal{E}(\phi) > \infty \), which by definition means that \( \phi \) is in the space \( \mathcal{E}^1_P(\mathbb{R}^n) \) of all functions with finite energy. In particular, integrating along affine lines in \( P_+(\mathbb{R}^n) \) the functional \( \mathcal{E} \) may be written as the following energy type functional

\[
(5.8) \quad \mathcal{E}(\phi) := \int_0^1 (\phi - \phi P) MA(\phi_0(1 - t) + t\phi) dt,
\]

which after expansions and integration over \( t \) can be written as a mixed Monge-Ampère expression (anyway, we will not use this representation). We may now rewrite (5.7) as

\[
\mathcal{F}(u) = \mathcal{E}(\Pi_X(\phi_0 + u))
\]

The following proposition is the key result in the variational approach to Monge-Ampère equations:

**Proposition 5.4.** [9] The functional \( \mathcal{F}(u) \) is Gateaux differentiable on \( C_b(\mathbb{R}^n) \) and

\[
d\mathcal{F}_u = MA(\Pi_X(\phi_0 + u))
\]

Similarly, if \( \phi_0 \) has finite energy then the corresponding statement also holds.

**Remark 5.5.** In the present real setting the previous proposition can be obtained from basic properties of the Legendre transform (see [9]), but it also holds in the more general complex setting (see Theorem B in [4]), where the proof is based on the complex analog of the following “orthogonality relation”

\[
\int_X MA(\Pi_X(\phi))(\Pi_X(\phi) - \phi) = 0,
\]

i.e. \( \Pi_X(\phi) = \phi \) almost everywhere with respect to \( MA(\Pi_X(\phi)) \).

Given a weighted set \((X, \phi_0)\) we next define the weighted energy \( E_{\phi_0}(\mu) \) as the Legendre transform of the functional \( \mathcal{F} \)

\[
E_{\phi_0}(\mu) := \sup_{u \in C_0^b(X)} (\mathcal{E}(\Pi_X(\phi_0 + u) - \int_X u\mu)
\]

i.e. by the formula (5.9) but replacing \( C^0(X) \) with the space \( C_b(X) \) of bounded continuous functions on \( X \) (recall that we are using a different sign convention than in the Legendre transform on \( \mathbb{R}^n \) defined by (5.1)).

**Proposition 5.6.** Let \( \mu \) be a probability measure on \( \mathbb{R}^n \) supported on a the closed set \( X \). Then

\[
(5.9) \quad E_{\phi_0}(\mu) := \sup_{\phi \in P_+(\mathbb{R}^n)} (\mathcal{E}(\phi) - \int (\phi - \phi_0)\mu)
\]

and \( E_{\phi_0}(\mu) \) is finite iff there exists \( \phi_\mu \in P(\mathbb{R}^n) \) such that

\[
(5.10) \quad E_{\phi_0}(\mu) = \mathcal{E}(\phi_\mu) - \int (\phi_\mu - \phi_0)\mu,
\]

which in turn is equivalent to \( \phi_\mu \) being a potential of \( \mu \) with finite energy, i.e. \( \phi_\mu \in \mathcal{E}^1_P(\mathbb{R}^n) \) is uniquely determined mod \( \mathbb{R} \) by the Monge-Ampère equation

\[
MA(\phi_\mu) = \mu
\]
Proof. To prove the first formula first observe that since \( \phi := \Pi_X(\phi_0 + u) \leq \phi_0 + u \) we immediately get the upper bound in (5.9). To get the lower bound we plug in the following function in the definition of \( E_{\phi_0}(\mu) : u := \phi_\mu - \phi_0 \), where \( MA(\phi_\mu) = \mu \) (compare below). By the domination principle for the Monge-Ampère operator (see the appendix) \( \Pi_X(\phi_\mu) = \phi_\mu \) and hence \( E(\phi_\mu) - \int(\phi_\mu - \phi_0)\mu \leq E_{\phi_0}(\mu) \). Finally, writing \( \phi_\mu \) as a decreasing sequence of elements in \( \mathcal{P}_+(\mathbb{R}^n) \) concludes the proof of formula (5.11) by basic continuity properties of \( E \) (see [9]). As for formula (5.10) it follows immediately from the variational construction of potentials in [9] (compare the proof of Prop 5.11 below).

We will often omit the subscript \( \phi_0 \) in the notation for the weighted energy. Note that, when \( \mu \) has compact support we can decompose

\[
(5.11) \quad E_{\phi_0}(\mu) = E_0(\mu) + \int \phi_0 \mu,
\]

where \( E_0(\mu) \) is independent of \( \phi_0 \). In fact, the formula applies to any \( \mu \) such that \( E_{\phi_0}(\mu) < \infty \) (for example, using an approximation argument; compare [8] for the complex setting).

The next proposition gives the relation to the theory of optimal transport:

**Proposition 5.7.** Let \( \mu \) be a probability measure on \( \mathbb{R}^n \). Then

\[
E_{\phi_0}(\mu) = C_{\phi_0}(\mu)
\]

where \( C_{\phi_0}(\mu) \) is the Monge-Kantorovich cost functional defined with respect to the target measure \( \lambda_P := 1dp \) on the target convex body \( P \) and the cost function \( c(x,p) = -x \cdot p + \phi_0(x) \) (see section 8.0.1).

**Proof.** By definition,

\[
E(\mu) = \sup_{u \in C_b(X)} \left( \int_P -\Pi_X(\phi_0 + u)^*(p)dp - \int u\mu \right)
\]

Now set \( v := (\Pi_X(\phi_0 + u))^* \) which is a bounded function (as explained in section 5.1.1). Using the extremal property of the Legendre transform we may rewrite the previous line as

\[
E(\mu) = \sup_{v \in C_b(P), u \in C_b(\mathbb{R}^n)} \left( \int_P -vdp - \int u\mu \right),
\]

where the sup ranges over all \( u \) and \( v \) such that \( -v - u \leq c(x,p) \), where \( c(x,p) = -x \cdot p + \phi_0(x) \). According to the general Kantorovich duality theorem [39] this means that \( E_{\phi_0}(\mu) = C_{\phi_0}(\mu) \) if the following condition on the cost function is satisfied: \( -c(x,y) \leq f(x) + g(p) \) for some functions \( f \) and \( g \) such that \( f \in L^1(X,\mu) \) and \( g \in L^1(P,dp) \). In the present setting we have \( -c(x,p) \leq \Pi_X(\phi_0)^* \) on \( X \times P \), where we recall that \( \Pi_X(\phi_0)^* \) is bounded on \( P \). Hence we may take \( f(x) = 0 \) and \( g(p) = \Pi_X(\phi_0)^* \) which thus concludes the proof of the alternative formula.

Given a weighed measure \( (\mu_0,\phi_0) \) we now define, following the notation in previous sections, the corresponding free energy functional on \( \mathcal{M}_1(X) \), where \( X \) is the support of \( \mu_0 \), by

\[
(5.12) \quad F_\beta = E_{\phi_0} + \frac{1}{\beta} D_{\mu_0},
\]

for any given \( \beta \in [0,\infty] \).
Proposition 5.8. The free energy functional $F_\beta$ defines a good rate functional on $\mathcal{M}_1(X)$.

Proof. Both functionals $E$ and $D$ are lsc, since they may be realized as Legendre transforms and to show that $F_\beta$ is a good rate functional on $\mathcal{M}(X)$ we thus only need to verify its properness (which is automatic when $X$ is compact). This could be proven directly, but anyway it is a general fact that a functional is a good rate functional if it is the rate functional of a LDP for which exponential tightness holds (see Lemma 1.2.18 in [24]) and the latter properties will be established in the proof of Theorem 1.1. \[\square\]

5.4. Proof of Theorem 1.1. The case when $\beta = \infty$ follows immediately from Prop 5.3 and the Gärtner-Ellis theorem (compare the proof in the complex setting considered in [8]). We hence consider the case when $\beta < \infty$ and to simplify the exposition we assume that $\beta_N/k = \beta$, but the proof in the general case is essentially the same.

Let us first consider the case when $X$ is compact. Setting $\beta_N = k$ Proposition 5.3 then allows us to apply the general Theorem 1.1 to deduce the LDP in Theorem 1.1. The fact that the potential of the minimizer $\mu_\beta$ solves the Monge-Ampère equation 1.3 is proved in section 5.1. But it can, for $X$ compact, also be seen as a special case of Theorem 5.6. Indeed, the latter results translates into $\mu = MA(\Pi_X u)$, for some continuous function $u$ on $X$ satisfying $MA(\Pi_X(\phi_0 + u)) = e^{\beta u} \mu_0$, where by the limiting construction used in the proof of 3.4 $\phi_0 + u$ is the restriction to $X$ of a function in the space $P(\mathbb{R}^n)$. But then it follows immediately that $\Pi_X(\phi_0 + u) = \phi_0 + u$ on $X$ and hence $\phi := \Pi_X(\phi_0 + u)$ solves the Monge-Ampère equation 1.3 and $\mu_\beta = MA(\phi)$, as desired.

In the general non-compact case we can use a variant of the localization argument used in the proof of Prop 5.3 to verify that the tightness assumption is satisfied and to reduce the problem to the compact set $B_R \cap X$. Indeed, let us first check the validity of the analog of Theorem 1.1. For the upper bound we still get, just as before,

$$E^{(N)}(\mu^{(N)}) \leq \inf_{\mu \in \mathcal{M}_1(\mathbb{R}^n)_c} (E(\mu) + \frac{1}{\beta} D(\mu)),$$

where $\mathcal{M}_1(\mathbb{R}^n)_c$ denotes the space of all compactly supported probability measures on $\mathbb{R}^n$. Anyway, by a simple approximation argument, where $\mu$ gets replaced with $1_{B_R} \mu/\mu(B_R)$ for a sequence of $R \to \infty$ we may as well take the infimum appearing in the right hand side above over all of $\mathcal{M}_1(\mathbb{R}^n)$ (see [8] for the complex case).

To prove the lower bound we first observe that the first marginals $\mu_1^{(N)}$ define a tight sequence. Indeed, by the inequality in Lemma 5.2 we have that

$$\mu_1^{(N)} \leq C e^{\beta(\pi_X(\phi) - \phi_0))} \mu_0 \leq C' e^{\beta(\phi - \phi_0)} \mu_0.$$

By assumption, the measure appearing in the rhs above has finite total mass on $\mathbb{R}^n$ and hence the tightness of the sequence $\mu_1^{(N)}$ follows. But then it follows from the standard weak pre-compactness of tight sequences that $\mu_1^{(N)}$ has a limit point $\mu_*$ in $\mathcal{M}_1(\mathbb{R}^n)$. Moreover, repeating the argument for the lower bound in the proof of Theorem 1.1 and using the asymptotics in 5.3 for the infimum of $(H_{\phi_0}^{(N)} + u)$, for any $u \in C_b(\mathbb{R}^n)$, gives

$$E(\mu_*) + D(\mu_*)/\beta \leq E^{(N)}(\mu_1^{(N)}) \leq \inf_{\mu \in \mathcal{M}_1(\mathbb{R}^n)} (E(\mu) + D(\mu)/\beta),$$
But then it follows from Prop [5.11] that \( \mu_* \) coincides with the unique minimizer of \( E(\mu) + D(\mu)/\beta \) and that its potential satisfies the desired equation.

Finally, to prove the LDP in the non-compact case we just have to verify the exponential tightness of the corresponding sequence \( \Gamma_k \). More precisely, that we need to prove is the space \( \mathcal{M}_1(X) \) may be exhausted by compact subsets \( \mathcal{F}_\alpha \) for \( \alpha > 0 \) such that \( \lim_{k \to \infty} \log(\Gamma_k(\mathcal{M}_1(X) - \mathcal{F}_\alpha)) / \beta N_k < -\alpha \). To prove this we let \( \mathcal{F}_\alpha \) be the set of all measures \( \mu \) on \( \mathcal{M}_1(X) \) such that \( \int (\phi_0 - \Pi_X \phi_0) \mu \leq 3 \alpha \). Since, by assumption \( \phi_0 - \Pi_X \phi_0 \to 0 \) at infinity in \( \mathbb{R}^n \), the set \( \mathcal{F}_\alpha \) is indeed compact. By definition

\[
\Gamma_k(\mathcal{M}_1(X) - \mathcal{F}_\alpha) = \frac{1}{Z_N} \int_{\{\phi_0 - \Pi_X \phi_0 > 3 \alpha N\}} \text{Per}_k(x_1, \ldots, x_N)^{\beta/k} \mu_0^{\otimes N_k}
\]

Now, applying Lemma [5.3] \( N \) times (i.e. one “coordinate at time”\( ) \) the density in the previous integral may be estimated from above by \( C_N^k e^{\epsilon N_k \beta} e^{-\beta(\phi_0 - \Pi_p \phi_0)} \) for some fixed small \( \epsilon > 0 \) (taken so that \( \epsilon < \alpha/2 \)). Hence, decomposing

\[
e^{-\beta(\phi_0 - \Pi_X \phi_0)} = e^{-\frac{1}{2} \beta(\phi_0 - \Pi_X \phi_0)} e^{-\frac{1}{2} \beta(\phi_0 - \Pi_X \phi_0)} \leq e^{-\frac{1}{2} \beta N_3 \alpha} e^{-\frac{1}{2} \beta(\phi_0 - \phi_p)} C^\beta
\]

and integrating wrt \( \mu_0^{\otimes N_k} \) (and using that \( e^{-\beta(\phi_0 - \phi_p)} \mu_0 \) is assumed to have finite total mass wrt any positive number \( \beta_* \)) finishes the proof of the exponential tightness.

**Remark 5.9**. In the case when \( X \) is compact we can directly combine Proposition [5.3] with the Gärtner-Ellis theorem to deduce that \( e^{-kH_N} \mu_0^{\otimes N} \) satisfies a LDP principle with speed \( kN \) and rate functional \( E(\mu) \). Then, arguing as in section [4.6] (and using that, by Prop [5.7] \( E = C \)) it follows that

\[
\sup_{X^N} \left| -\frac{1}{k} \log \text{Per} (x_1, \ldots, x_N) - C \left( \frac{1}{N}(\delta_{x_1} + \cdots + \delta_{x_N}) \right) \right| = 0,
\]

where \( C \) denotes the Monge-Kantorovich total cost functional associated to the cost function \( c(x, p) = -\langle x, p \rangle \).

**5.4.1. Proof of Cor [5.2]**. By Theorem [1.1] the corresponding one-point correlation measures \( \rho^{(N)}_1 \mu_0 \) converge weakly to \( \mu_\beta \). Since, the latter measure may be written as \( \mu_\beta = e^{\beta (\phi - \phi_0)} \mu_0 \) for \( \phi \) the unique solution in \( \mathcal{P}(\mathbb{R}^n) \) to equation [1.3] this means that

\[
\rho^{(N)}_1 \mu_0 \to e^{\beta (\phi - \phi_0)} \mu_0
\]

weakly. Now \( \phi^{(N)} := \frac{1}{N} \log \rho^{(N)}_1 - \phi_0 \) is a sequence in \( \mathcal{P}(\mathbb{R}^n) \) such that \( \int e^{\beta \phi^{(N)}} \mu_0 = 1 \) and hence, by the inequality in Lemma [5.2] fixing a point \( x_0 \) in \( X \) gives that \( \phi^{(N)}(x_0) \) is uniformly bounded from above. Hence, by Prop [5.1] we may, after passing to a subsequence, assume that \( \phi^{(N)} \to \phi_* \) locally uniformly for some function \( \phi_* \) in \( \mathcal{P}(\mathbb{R}^n) \). But then it follows from the convergence above that \( \phi_* = \phi \) almost everywhere wrt \( \mu_0 \) and hence everywhere if the support of \( X \) is all of \( \mathbb{R}^n \). Since, we may repeat the same argument for any subsequence of \( \phi^{(N)} \) it follows from the uniqueness of the solution \( \phi \) that the whole sequence \( \phi^{(N)} \) converges to \( \phi \), as desired.

**5.4.2. Proof of Cor [5.3]**. By assumption the support \( X \) of \( \mu_0 := \rho_0 1_X dx \) is compact and it is the closure of its interior. First note that \( -\phi^{(N)}(x) \) represents the differential of \( E_N(\mu_0) := \int_{X^N} H^{(N)}(\mu_0)^{\otimes N} \). More precisely, if we fix a smooth probability density \( \rho \) with compact support in the interior of \( X \) and set \( \mu_t := \mu_0 + t(\rho - \rho_0)dx \),
for $t$ sufficiently small, i.e. $|t| \leq \epsilon$ such that $\mu_t \in M_1(X)$, then a direct calculation gives $dE_N(\mu_t)/dt|_{t=0} = -\int_X \phi^{(N)}(x)(\rho - \rho_0)dx$. Next, we note that, since $X$ is compact $E(\mu_t) < \infty$, where $E(\mu)$ denotes the unweighted energy (i.e. $\phi_0 = 0$) and $\lim_{N \to \infty} E_N(\mu_t) = E(\mu_t)$ (by Theorem 2.1). Hence, by Lemma 4.5 and the uniqueness mod $\mathbb{R}$ of potentials it follows that $E(\mu_t)$ is differentiable and $dE(\mu_t)/dt|_{t=0} = -\int_X \phi^{(N)}(x)(\rho - \rho_0)dx$, where $\phi^{(N)}$ is any potential for $\mu_t$ which we may as well take to be the one uniquely determined by the normalization condition $\int \phi^{(N)} \mu_0 = 0$. Since $E(\mu_t)$ is moreover convex it then follows from basic properties of convex functions that $\lim_{N \to \infty} dE_N(\mu_t)/dt = dE(\mu_t)/dt$ (compare the proof of Lemma 2.4), i.e.

$$\lim_{N \to \infty} \int_X \phi^{(N)}(x)(\rho - \rho_0)dx = \int_X \phi^{\mu_0}(x)(\rho - \rho_0)dx$$

(5.13)

Now, by assumption, $\int \phi^{(N)} \mu_0 = 0$ and thus we may, by Prop 5.4, after perhaps passing to a subsequence, assume that $\phi^{(N)} \to \phi_*$ in $\mathcal{P}(\mathbb{R}^n)$ such that $\int \phi_* \mu_0 = 0$. But since $\phi^{\mu_0}$ satisfies the same normalization condition and (5.13) holds for any $\rho$ as above we conclude that $\phi_* = \phi^{\mu_0}$ on the interior of $X$. In particular, $MA(\phi_*) = \mu_0$ on the interior of $X$ and since $\int_{\mathbb{R}^n} MA(\phi_*) \leq 1$ and $\int_{\mathbb{R}^n} \mu_0 = 1$ it follows that $MA(\phi_*) = \mu_0$ on all of $\mathbb{R}^n$ and hence, by uniqueness of normalized potentials, $\phi_* = \phi^{\mu_0}$ on all of $\mathbb{R}^n$, as desired.

Next, it follows form the regularity result in [23] that $\phi$ is $C^{1,\alpha}$-smooth in the interior of $X$ for some $\alpha > 0$. We briefly recall the argument: first the solution $\phi$ in $\mathcal{P}(\mathbb{R}^n)$ of the Monge-Ampère equation $MA(\phi) = \mu_0$ (in the sense of Alexandrov) has, by assumption, a Monge-Ampère measure $MA(\phi)$ which is absolutely continuous wrt Lebesgue measure and vanishes on the complement of $X$. Hence, its restriction to the interior of $X$ defines a Brenier solution on $X$, i.e. it satisfies the MA-equation on the interior of $X$ in the weak sense of Brenier and the almost everywhere defined map $\nabla \phi$ from the interior of $X$ to $P$ is almost everywhere surjective. But as shown in [23] any Brenier solution $\phi$ (which is in fact uniquely determined) is $C^{1,\alpha}$-smooth in the interior of $X$ (the point is that, as shown in [23], it is strictly convex and then the regularity follows from [21]). Finally, since $\phi$ is convex and differentiable on the interior of $X$ the previous convergence of $\phi^{(N)}$ implies, by basic properties of convex functions, point-wise convergence everywhere on the interior of $X$ for the corresponding gradients.

5.5. Variational properties of Monge-Ampère equations and regularity. In this section we will establish some properties of the Monge-Ampère equations and convex envelopes studied above, most of which can be reduced to essentially known results.

**Proposition 5.10.** Let $(X, \phi_0)$ be a weighted set and denote by $\phi_c := \Pi_X \phi_0$ the corresponding convex envelope. Then

- \( \mu_c := MA(\phi_c) \) is the unique minimizer of the functional $\mu \mapsto E_{\phi_0}(\mu)$ on $M_1(X)$.
- If $X = \mathbb{R}^n$ and $\phi_0$ is smooth, then $\phi_c$ is locally $C^{1,1}$, i.e. $\nabla \phi_c$ is locally a Lipschitz map and
  \[
  \mu_c = 1_{D_0} \det(\frac{\partial^2 \phi}{\partial x_i \partial x_j})dx.
  \]
where $D_0$ is the compact set where $\phi_c = \phi_0$ (and the density above is pointwise defined a.e. on $D_0$).

**Proof.** The minimizing property: By construction $E_{\phi_0} (\mu)$ is the Legendre transform of the functional $F_X (u) := \mathcal{E} (\Pi_X (\phi_0 + u))$ and hence, since the latter functional is Gateaux differentiable, it follows from general properties of Legendre transforms (Lemma 1.2) that $E_{\phi_0} (\mu)$ admits a unique minimizer, which is given by the differential $d\mathcal{E} (\Pi_X (\phi_0 + u))$ at $u = 0$. Finally, by Prop 5.4 the latter differential is equal to $MA (\Pi_X \phi_0)$.

**Regularity:** This is a special case of the regularity results for the generalized Lelong class obtained in [2], modeled on the classical approach of Bedford-Taylor. □

**Proposition 5.11.** Let $\mu_0$ be a measure on $\mathbb{R}^n$ and $\phi_0$ a continuous function on $\mathbb{R}^n$ such that $\phi_0 - \phi_F$ is proper and assume that $\int e^{\beta (\phi_F - \phi_0)} \mu_0 < \infty$. Then

1. The Monge-Ampère equation (3.2) admits a solution $\phi_\beta$ in $\mathcal{P} (\mathbb{R}^n)$ of finite energy, i.e. $\mathcal{E} (\phi_\beta) > -\infty$.
2. Any two solutions of full Monge-Ampère mass coincide up to an additive constant
3. The probability measure $\mu_\beta := MA (\phi_\beta)$ is the unique minimizer of the corresponding free energy functional $F_\beta$.
4. If $\phi_0$ is smooth and $\mu_0 = \rho_0 dx$ for a strictly positive smooth function $\rho_0$, then the solution $\phi_\beta$ is also smooth.

**Proof.** Existence: The existence of a weak solution is well-known, at least when $\mu_0$ is absolutely continuous wrt $dx$ [1], but it may be illuminating to give a variational proof of the general case, in the spirit of the present paper. The point is that, following the variational approach in [1], we just need to verify that the following coercivity estimate holds:

$$D (\phi) := -\mathcal{E} (\phi) + \frac{1}{\beta} \log \int e^{\beta (\phi_F - \phi_0)} \mu_0 \geq -\mathcal{E} (\phi) + (\phi - \phi_0) (x_0)$$

where $x_0$ is a fixed point in the support $X$ of $\mu_0$. But this follows immediately from the inequality

$$\sup_X e^{\beta (\phi - \phi_0)} \leq C \int_{X \cap B_R} e^{\beta (\phi - \phi_0)} \mu_0,$$

which is a direct consequence of the inequality in Lemma 6.2. With the coercivity inequality in place the solution $\phi$ may be obtained as a minimizer of the functional $D$. Indeed, since $D (\phi + c) = D (\phi)$ we can take a sequence of functions $\phi_j \in \mathcal{P}_+ (\mathbb{R}^n)$ such that $D (\phi_j)$ converges to the supremum of $D$ on $\mathcal{P} (\mathbb{R}^n)$ and such that $\phi_j$ is normalized, i.e. $(\phi_j - \phi_0) (x_0) = 0$. By the Arzelà-Ascoli theorem $\phi_j$ converges, after perhaps passing to a subsequence, locally uniformly to $\phi_\infty$ in $\mathcal{P} (\mathbb{R}^n)$. Moreover, by the coercivity estimate above $\phi_\infty$ has finite energy. Defining

$$\tilde{D} (\phi) := -\mathcal{E} (\Pi_X \phi) + \frac{1}{\beta} \log \int e^{\beta (\phi - \phi_0)} \mu_0$$

we have $\tilde{D} (\phi) \geq \tilde{D} (\Pi_X \phi)$ and hence, if $u$ is any given smooth compactly supported function on $\mathbb{R}^n$ then the function $t \mapsto \tilde{D} (\phi + tu)$ on $\mathbb{R}$ has a maximum at $t = 0$ and is differentiable by Prop 5.4. Hence, its derivative at $t = 0$ vanishes and the formula for the differential in Prop 5.4 then shows that $\phi_\infty$ satisfies the desired equation up
to a multiplicative normalization factor which can be removed by adding a constant to $\phi_\infty$.

**Uniqueness:** The uniqueness of finite energy solutions can be shown by convexity arguments, but, anyway, the general case follows form the comparison principle for $MA$. Indeed, if $u$ and $v$ are in $\mathcal{E}(X)$ then the comparison principle says

$$\int_{\{u<v\}} MA(v) \leq \int_{\{u<v\}} MA(u)$$

But if $u$ and $v$ are solutions of equation [13] then it must be that $u = v$ a.e wrt the measure $\mu_0$ and hence $MA(u) = MA(v) = \mu_0$, which implies that $u - v$ is constant, by the uniqueness of normalized potentials of any probability measure.

**Minimizing property:** This is proved precisely as in the proof of Theorem 3.6

**Regularity:** This is proved exactly as in [9], using Caffarelli’s interior regularity results. Briefly, since $\phi$ has finite energy the image of the corresponding subgradient map to $P$ is surjective in the almost everywhere sense. But this implies that $\phi$ is proper (using that we may assume that 0 is in an interior point in $P$), i.e. the sublevel sets $\Omega_R := \{ \phi < R \}$ are bounded convex domains exhausting $\mathbb{R}^n$. On $\Omega_R$ the function $u := \phi - R$ defines a function in $C_0(\Omega_R)$, vanishing on the boundary to which we may apply the regularity results in [21, 23, 20, 22] to deduce that $u$ is smooth (see [29] for the complete argument).

**Remark 5.12.** Note that if there is a solution $\phi$ in $\mathcal{P}_+ (\mathbb{R}^n)$ then necessarily $\int e^{\beta (\phi_P - \phi_0)} \mu_0 \leq C \int e^{\beta (\phi - \phi_0)} \mu_0 = C \int MA(\phi) \leq C < \infty$.

**Proposition 5.13.** Let $(\mu_0, \phi_0)$ be a weighted measure and let $X$ be the support of $\mu_0$. Denote by $\phi_\beta$ be the unique solution in in $\mathcal{P} (\mathbb{R}^n)$ to the corresponding equation [13]. Then $\phi_\beta$ converges, as $\beta \to \infty$, locally uniformly to the envelope $\Pi_X \phi_0 := \phi_e$ iff $\mu_e$ has finite entropy with respect to $\mu_0$. In particular, this is the case if $X = \mathbb{R}^n$ and $\phi_0$ is smooth.

**Proof.** Let us first verify that the family $\mu_\beta := MA(\phi_\beta)$ is tight. By the inequality in Lemma 5.2 we have, since $\int e^{\beta (\phi_\beta - \phi_0)} \mu_0 = 1$, that

$$\int_{\{x \in X - B_R\}} (\phi_\beta - \phi_0) \leq C/\beta + \phi_e - \phi_0 \leq C/\beta + C' + \phi_P - \phi_0.$$ 

Now, by assumption, $\phi_P - \phi_0 \to -\infty$ and hence there exists $\delta > 0$ such that

$$x \in X - B_R \implies (\phi_\beta - \phi_0) \leq (1 - \delta)(\phi_P - \phi_0)$$

for some large ball $B_R$ (where we may assume that $(\phi_P - \phi_0) < 0$). But then

$$\int_{X - B_R} e^{\beta (\phi_\beta - \phi)} \mu_0 \leq \int_{X - B_R} e^{\beta(1 - \delta)(\phi_P - \phi)} \mu_0 = \epsilon_R,$$ 

where $\epsilon_R \to 0$, as $R \to \infty$ (since, by assumption, the integral is finite for some $\beta$). Thus the family $\mu_\beta$ is tight as desired. In particular, the family admits a limit point $\mu_\infty$ in $\mathcal{M}_1(\mathcal{X})$ and we next show that it coincides with $\mu_e$, the unique minimizer of the energy functional $E$ on $\mathcal{M}_1(\mathcal{X})$. This can be proved by following the argument given in the complex case in [8] (Theorem 3.13). But here we note that a simpler argument can be given in the real setting. Let us first assume that $\mu_0(\mu_e) < \infty$. Since $\mu_\beta$ minimizes the functional $F_\beta$ we then get

$$E(\mu_e) = \lim_{\beta \to \infty} F_\beta(\mu_e) \geq \limsup_{\beta \to \infty} F_\beta(\mu_\beta) = E(\mu_\infty),$$

using that $E$ is lower semi-continuous and that $D_{\mu_\beta}(\mu_\beta) \leq C$ (by [13]) to get that last inequality. Hence, $\mu_\infty = \mu_e = MA(\phi_e)$ by Prop 5.10. In other words
\[ MA(\phi_\beta) \to MA(\phi_c) \] weakly. But after passing to a subsequence we may assume that \( \phi_\beta \to \phi_\infty \) for some \( \phi_\infty \in \mathcal{P}(\mathbb{R}^n) \) and since \( \phi_\beta \) has finite energy and hence full Monge-Ampère mass it follows that \( MA(\phi_\beta) \to MA(\phi_\infty) = \mu_c \) (compare [9]). But by the uniqueness mod \( \mathbb{R} \) of potentials it then follows that \( \phi_\infty = \phi_c + C \). Finally, to see that \( C = 0 \) we set \( u := \phi_\infty - \phi_0 \) which is continuous and bounded from above. Hence, by the equations for \( \phi_\beta \) we have that \( 0 = \lim_{\beta \to \infty} (\log \int e^{\beta u} \mu_0) / \beta = \sup_X u \), which forces \( C = 0 \), using that \( \sup_X (\Pi_X \phi_0 - \phi_0) = 0 \) (indeed, the incidence set \( D_{\phi_0} \) is non-zero, since it contains the support of the probability measure \( MA(\Pi_X \phi_0) \)).

To get the converse statement we assume that \( \phi_\beta \to \phi_c \). But then \( \mu_\beta \to \mu_c \) and since \( D_{\mu_0} \) is lsc and \( D_{\mu_0}(\mu_\beta) \leq C \) it thus follows that \( D_{\mu_0}(\mu_c) < \infty \). \( \square \)

6. General target measures and random allocation of the target points

The proof of Theorem 1.1 in fact applies to a more general setting where \( -x \cdot p \) is replaced by a function \( c(x, p) \) and \( \lambda_P \) is replaced with a probability measure \( \nu \) on \( \mathbb{R}^n \). One furthermore needs to fix a sequence \( \beta_N^* \) (playing the role of \( k \)) such that

\[ \beta_N^* \to \infty, \quad \frac{\log N!}{N \beta_N^*} \to \infty \]

and a sequence of \( N \)-tuples \( p^{(N)} := (p_1^{(N)}, \ldots, p_N^{(N)}) \) such that

\[ \frac{1}{N} \sum_{i=1}^N \delta_{p_i^{(N)}} \to \nu, \]

weakly as \( N \to \infty \). Given this data one then replaces \( \text{Per}(x_1, \ldots, x_N) \) with the permanent

\[ \text{Per}_c(x_1, \ldots, x_N) := \text{Per}(e^{-\beta N c(x_i, p_i)})_{i,j \leq N} \]

and sets

\[ \mu_{\beta_N}^{(N)} := \frac{\text{Per}_c(x_1, \ldots, x_N)^{\beta N \beta_N^*}}{Z_{N, \beta_N^*}^{\beta N}} \mu_0^{\otimes N} \]

Under suitable regularity assumptions on \( c(x, y) \) and \( \nu \) the previous proof of Theorem 1.1 generalize verbatim to this more general setting (for example, the proof applies if \( c(x, p) \) is continuous and uniformly Lipschitz wrt \( x \) as \( p \) ranges over the support of \( \nu \) and \( \nu \) is absolutely continuous with respect to \( \lambda_P \)). The deterministic measure \( \mu_\beta \) appearing in the limit then coincides with the minimizer of the corresponding free energy functional defined with respect to the cost functional \( C(\mu, \nu) \) (but for a general cost \( c(x, p) \) it can not be directly linked to a Monge-Ampère equation). The key point is that the proof of Prop 5.3 still applies if one replaces the Legendre transform \( \phi^*(p) \) with \( \phi^* c(x, p) \) and the measure \( \lambda_P \) with \( \nu \) (up to signs this is the same transform as the one appearing in [26]). Then the Kantorovich duality argument used in the proof of Prop 5.7 shows that the corresponding functional \( E(\mu) \) coincides with the optimal cost functional \( C(\mu, \nu) \), determined by \( c(x, p) \) (see the appendix).

6.1. Random allocation of target points and quenched variables. Let us in particular consider the case when we still have \( c(x, p) = -x \cdot p \), but replacing \( \lambda_P \) with a general target measure \( \nu \) absolutely continuous with respect to \( \lambda_P \), where, as before, \( P \) denotes a given convex body. As before we also assume given a measure \( \mu_0 \) on \( \mathbb{R}^n \) and for simplicity we will assume that its support \( X \) is compact and that the
Then the following weak convergence of measures on $X$ holds:

$$\nu \rightarrow \nu_0.$$ 

To the data $(\mu_0, \nu, \beta)$ we may now associate the following Monge-Ampère equation:

$$MA_{\nu}(\phi) = e^{\beta\phi} \mu_0,$$

assuming as before that $\nabla \phi$ maps $\mathbb{R}^n$ into $P$. If $\nu = 1_p e^{-\psi_0(p)} dp$ and $\phi$ is smooth the previous equation just means that

$$MA(\phi) e^{-\psi_0(\nabla \phi)} = e^{\beta \phi} \mu_0.$$

In particular, in the case $\beta = 0$ the corresponding equation appears in the optimal transport problem defined with respect to the target measure $\nu$. We will denote the corresponding optimal cost function by $C(\mu, \nu)$, which, as before, will be considered as a functional of $\mu$. If one would also fix a sequence of $p^{(N)}$ approximating $\nu$ in the sense of [6.1] then, as explained above, the previous results apply to this more general setting. However, it is also interesting to see that there is a variant of this setting which does not depend on fixing a sequence of $p^{(N)}$ and to which we next turn. The idea is to view all the previously defined objects, such as $\mu$, $p$, etc as random variables on $(P^N, \nu^{\otimes N})$. This means that we view the variables $p_i$ appearing in

$$\text{Per}(x_1, ..., x_N, p_1, ..., p_N) := \text{Per}(e^{-\beta N \psi_0(x_i, p_i)})_{i,j \leq N}$$

as independent random variables identically distributed according to probability measure $\nu$. In other words we perform a random allocation of the $p_i$'s according to the measure $\nu$ (which is somewhat related to the setting considered in [29]). In the terminology appearing in the mathematics of disordered systems we thus view the variables $(p_1, ..., p_N)$ as quenched (i.e. frozen); compare [15].

We will denote by $E$ expectations defined with respect to the ensemble $(P^N, \nu^{\otimes N})$. The previous arguments can then be adapted to prove the following variant of Theorem 1.1 (or rather Theorem 4.3):

**Theorem 6.1.** Assume given data $(\mu_0, \nu, \beta)$ as above and a sequence $\beta_N \rightarrow \beta > 0$. Then the following weak convergence of measures on $X$ holds, as $N \rightarrow \infty$:

$$E(\int_{X^{N-1}} \mu^{(N)}_{\beta_N}) \rightarrow MA(\phi_\beta),$$

where $\phi_\beta$ is a solution to the equation (6.3).

Note that we may equivalently view $E(\int_{X^{N-1}} \mu^{(N)}_{\beta_N})$ as the expectation of the empirical measure $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ with respect to the following probability measure on $X^N \times P^N$:

$$\hat{\gamma}^{(N)}_{\beta_N} := \frac{\text{Per}(x_1, ..., x_N, p_1, ..., p_N)^{\beta_N/\beta_N^*}}{\int_{X^N} \text{Per}(x_1, ..., x_N, p_1, ..., p_N)^{\beta_N/\beta_N^*} \mu_0^{\otimes N} \otimes \nu^{\otimes N}} \mu_0^{\otimes (N-1)} \otimes \nu^{\otimes N}$$

The convergence of the expected one-point correlation measures in the previous theorem yields for any $\beta > 0$, just as before, a sequence of explicit approximate solutions $\phi_{N, \beta}$ to the real Monge-Ampère equations (6.3) (but now integrating wrt $\mu_0^{\otimes (N-1)} \otimes \nu^{\otimes N}$). For the case $\beta = 0$ we also obtain a variant of Cor 1.3 which may be formulated as follows:
Corollary 6.2. Assume given two probability measures $\mu_0$ and $\mu_1$ of the form $\mu_i = \rho_i 1_{X_i} \, dx_i$ such that $X_i$ is the closure of a bounded domain whose boundary $\partial X_i$ is a null set for Lebesgue measure and assume that $\rho_i$ is bounded from below and above by positive constants on $X$. Assume also that $X_1$ is convex. Set
\[
\phi_N(x_1) := \frac{1}{N} \int_{X^{N-1} \times P^N} \log \text{Per}(x_1, \ldots, x_N, p_1, \ldots, p_N) \mu_0^{N-1} \otimes \mu_1^N - c_N,
\]
where $c_N$ is the normalizing constant ensuring that $\int_{X^N} \phi(N) \mu_0 = 0$. Then $\phi_N$ converges, as $N \to \infty$, locally uniformly to the unique convex function $\phi$ solving the equation \( \partial_{\beta} \phi \) for $\beta = 0$ and such that $\nabla \phi$ maps $X_0$ almost surjectively onto $X_1$. Moreover, $T(N) = \nabla \phi(N)$ converges point-wise, in the interior of $X_0$, to the (Hölder continuous) optimal map $T$ for the Monge problem of transporting the probability measure $\mu_0$ to $\mu_1$, where the optimality is defined with respect to the cost function $c(x, p) = |x - p|^2$.

6.1.1. Proof of Theorem 6.1. As the arguments are similar to the previous ones, we will be rather brief. Let us first show that
\[
\lim_{N \to \infty} \mathbb{E}(-\frac{1}{N} \log Z_{N,\beta}) = \inf_{\mu \in \mathcal{M}_1(X)} (C(\mu, \nu) + D_{\mu_0}(\mu)/\beta)
\]
To this end we first observe that $-\frac{1}{N} \log Z_{N,\beta}$ viewed as a function of the quenched variables $(\rho_1, \ldots, \rho_N)$ is Lipschitz continuous in each coordinate $\rho_i$ with a uniform Lipschitz constant $L$ (which is proportional to the diameter of $X$). Indeed, this follows immediately from the fact that for any fixed $(x_1, \ldots, x_N)$ the Hamiltonian $H_N/N$ has the corresponding Lipschitz property, which in turn follows from the fact that $\partial c(x, p)/\partial p = x$ is uniformly bounded, since we have assumed that $X$ is compact. Moreover, by the same argument $-\frac{1}{N} \log Z_{N,\beta}$ is uniformly bounded. Now, using (a weak form of) Sanov’s theorem we may replace the integration over $P^N$ with the integral over a ball $B_\delta(\nu)$ of a fixed small radius $\delta$ centered at $\nu$ in the space $\mathcal{M}_1(P)$ of all probability measures on $P$. Then we pick a sequence $\rho^{(N)} \in P^N$ approximating $\nu$ in the sense of 6.1 and in particular $\delta^{(N)}(\rho^{(N)}) \in B_\delta(\nu)$ for $N$ sufficiently large. The point is that, as explained above, along this sequence we have
\[
\lim_{N \to \infty} (-\frac{1}{N} \log Z_{N,\beta}) = \inf_{\mu \in \mathcal{M}_1(X)} (C(\mu, \nu) + D_{\mu_0}(\mu)/\beta)
\]
Finally, by the uniform Lipschitz estimate and the uniform bound on $-\frac{1}{N} \log Z_{N,\beta}$ the oscillation of $-\frac{1}{N} \log Z_{N,\beta}$ on $B_\delta(\nu)$ is bounded by a uniform constant times $\delta$ and hence the previous convergence implies the convergence in 6.4 by first letting $N \to \infty$ and then $\delta \to 0$.

Now, fixing a continuous function $u$ on $X$ we can repeat the previous argument with $H^{(N)}$ replaced with $H^{(N)} + u$ to get
\[
\Lambda_N[u] := \mathbb{E}(-\frac{1}{N} \log Z_{N,\beta}[u]) \to \Lambda[u] := \inf_{\mu \in \mathcal{M}_1(X)} \left( C(\mu, \nu) + D_{\mu_0}(\mu)/\beta + \int u \mu \right)
\]
Next we observe that the measure $\mathbb{E}(\int_{X^{N-1}} \mu^{(N)}_{\beta})$ on $X$ represents the differential at $u = 0$ of the functional $\Lambda_N[u]$ on $C^0(X)$. But the latter functional is convex and converges to $\Lambda$ whose differential at 0 is the unique minimizer $\mu_\beta$ of the functional on $\mathcal{M}_1(X)$ appearing in the rhs above 6.5. But then it follows, as before, by
basic convex analysis that $\mu_\beta$ represents the differential of $A$ at $u = 0$ and that
$\mathbb{E}(\int_{X^{N-1}} \mu_{\beta N}^{(N)})$ converges to $\mu_\beta$, as desired.

6.2. **Proof of Cor 6.2** Switching the order of integration we can write

$$
\mathbb{E}(E^{(N)}(\mu)) = \frac{1}{N} \int_{X^N} \left( \int_{P^N} \frac{1}{\beta_N} \log \sum_{\sigma \in S_N} e^{\beta_N(x_1 p_{\sigma(1)} + \ldots + x_N p_{\sigma(N)})} \nu_{\gamma N} \right) \mu_{\gamma N}^N
$$

Now we can proceed precisely as in the proof of the previous theorem by localizing the integration over $P^N$ to a ball $B_\delta(\nu)$ of a fixed small radius $\delta$ centered at $\nu$ in the space $\mathcal{M}_1(P)$ of all probability measures on $P$ and picking a sequence $p^{(N)} \in P^N$ approximating $\nu$ in the sense of 6.1. The point is that, as explained above, along this sequence we have $E^{(N)}_{p^{(N)}}(\mu) \to C(\mu, \nu)$. Finally, by the Lipschitz uniform estimate for $H^{(N)}/N$ the oscillation of $H^{(N)}/N$ on $B_\delta(\nu)$ is bounded by a uniform constant $\delta$ and hence

$$
\lim_{N \to \infty} \mathbb{E}(E^{(N)}(\mu)) = \lim_{N \to \infty} E^{(N)}_{p^{(N)}}(\mu) = C(\mu, \nu)
$$

Now the proof is concluded by differentiating with respect to $\mu$, precisely as in the proof of Cor 1.3. Note that the convexity of $P$ is crucial to get the Hölder regularity of the transport map, as explained in [23].

7. **Outlook**

7.1. **Relation to the complex setting, determinantal point processes and toric varieties.**

7.1.1. **The toric setting.** In this section we come back to the original setting where $\nu = \lambda_\rho$. Consider the complex torus $\mathbb{C}^n$ and denote by $T^n$ the corresponding real unit-torus in $\mathbb{C}^n$, which acts, in the standard way, holomorphically on $\mathbb{C}^n$. Denote by $Log$ be the map

$$
Log : \mathbb{C}^n \to \mathbb{R}^n
$$

from $\mathbb{C}^n$ to $\mathbb{R}^n$ defined by $x := Log(z) := x$, where $x$ is the vector whose $j$ th coordinate is the log of the squared absolute value of the $j$ th coordinate of $z$. The fibers are thus the orbits of the real torus $T^n$ on $\mathbb{C}^n$. The definition is made so that, if $p \in \mathbb{Z}^n$, then $|z|^p = e^{p \cdot x}$ in multiindex notation.

Denote by $\Delta^{(N)}(z_1, \ldots, z_N)$ the Vandermonde determinant on $(\mathbb{C}^n)^N$ determined by the convex body $P$, i.e.

$$
\Delta^{(N)}(z_1, \ldots, z_N) = \det(z_i^{p_j}),
$$

where $p_j$ ranges over the $N$ lattice points in $kP$. Fixing a measure $\tilde{\mu}_0$ on $\mathbb{C}^n$ and a continuous function $\tilde{\phi}_0(z)$ of suitable growth one obtains, for any sequence $\beta_N$ of positive numbers, a probability measure $\tilde{\mu}_N^{(N)}$ on $(\mathbb{C}^n)^N$ by setting

$$
\tilde{\mu}_{\beta_N}^{(N)} := \frac{1}{Z_{N, \beta_N}} |\Delta^{(N)}(z_1, \ldots, z_N)|^{2\beta_N/k} e^{-k(\tilde{\phi}_0(z_1) + \ldots + \tilde{\phi}_0(z_N))} \tilde{\mu}_0^N
$$

For $\beta_N = k$ this is defines a determinantal point process (see [28] for general properties of such processes and [6] for large deviation results for these particular determinantal processes). The relation to the present paper stems from the simple observation that if the background data $(\tilde{\mu}_0, \tilde{\phi}_0)$ is $T^n$-invariant, then the push-forward of the corresponding determinantal point process is a permanental point process. More precisely, in the case $\beta_N = k$, the push-forward of the corresponding
probability measure \( \tilde{\mu}^{(N)} \) is precisely the permanental probability measure \( \tilde{\mu}^{(N)} \) studied in the present paper, determined by the weighted measure \((\mu_0, \phi_0)\), where \( \phi_0 = \log^* \phi_0 \) and \( \log_\ast \mu_0 = \mu_0 \) (i.e. abusing notation slightly \( \phi(x) = \phi(z) \) and \( \tilde{\mu} = \mu \wedge d\theta \), where \( d\theta \) denotes the invariant probability measure on \( T^n \)). To see this just note that expanding \( \Delta^{(N)}(z_1, ..., z_N) \) as an alternating sum over the permutations \( \sigma \) in \( S_N \) and using Parseval’s formula for \( x \) fixed to carry out the integration over the corresponding torus fiber \( \log^{-1}(\{x\}) \) immediately gives

\[
(\log)^\ast \left( \Delta^{(N)}(z_1, ..., z_N) |(\mu_0 \otimes d\theta)^{\otimes N} \right) = \text{Per}(x_1, ..., x_N)(\mu_0)^{\otimes N}.
\]

Moreover, it is also interesting to see that there is a pluripotential analog of this determinantal/permanental correspondence: denoting by \( E(\tilde{\mu}) \) the pluricomplex energy of a probability measure \( \tilde{\mu} \) on \( \mathbb{C}^n \) (defined with respect to the reference weight \( \log^* \phi_P \) one gets, if \( \tilde{\mu} \) is \( T^n \)-invariant, that

\[
E(\tilde{\mu}) = C(\mu),
\]

where, as before, \( C(\mu) \) is the optimal cost functional functional defined with respect to the cost function \( c(x, p) := -x \cdot p \) and the target measure \( \lambda_P \). This follows immediately from Prop \([4,7]\) combined with the essentially well-known fact that, when \( \mu \) is \( T^n \)-invariant, \( E(\mu) \) can be expressed in terms of the Legendre transform of the Monge-Ampère potential of \( \mu \). The key point is the basic fact that if \( \phi \) is a \( T^n \)-invariant plurisubharmonic function (i.e. \( \partial \bar{\partial} \phi \geq 0 \)) then \( \phi \) is convex and

\[
\log^\ast (\text{MA}_\mathbb{C}(\phi)) = \text{MA}(\phi),
\]

where \( \text{MA}_\mathbb{C} \) denotes the \emph{complex} Monge-Ampère operator, i.e.

\[
\text{MA}_\mathbb{C}(\psi) := \left( \frac{i}{2\pi} \partial \bar{\partial} \psi \right)^n / n! \left( = c_n \det \left( \frac{\partial \psi}{\partial z_i \partial \bar{z}_j} \right) \right).
\]

Alternatively, the relation \([3,8]\) follows from the correspondence \([2,2]\) by combining the large deviation principle for the determinantal point processes in \([6]\), where \( E \) appears as the rate functional, applied to the toric case, with the large deviation principle for the corresponding permanental point-process proved in the present paper. Strictly speaking the results in \([6]\) only apply when \( P \) is a rational polytope, but the proofs are essentially the same in the general case (compare the setting in \([2]\) ). The point is that when \( P \) is a rational polytope it defines a toric variety \( X_P \) with an ample line bundle \( L_P \) to which the results in \([6]\) can be applied. Briefly, the toric variety \( X_P \), which is an equivariant compactification of \( \mathbb{C}^n \), may be defined as the projective algebraic variety obtained as the closure in complex projective space \( \mathbb{P}^N \) of the affine algebraic variety in \( \mathbb{C}^N \) defined by the image of the map

\[
\mathbb{C}^n \to \mathbb{C}^N, \quad z \mapsto (z^{P_1}, ..., z^{P_N}),
\]

for \( k \) sufficiently large and \( L_P \) is the restriction of the hyperplane line bundle on \( \mathbb{P}^N \) (see \([3]\) and references therein).

It should be stressed that, unless \( \beta_N = k \), the push-forward under the map \( \log \) of the probability measure \( \tilde{\mu}^{(N)}_{\beta_N}(\mathbb{C}^n)^N \) is \emph{not} equal to the corresponding probability measure \( \mu^{(N)}_{\beta_N}(\mathbb{R}^n)^N \). Still, one would expect that this is true in an asymptotic sense, as \( N \to \infty \).
7.1.2. The general complex geometric setting and Kähler-Einstein geometry. The general complex geometric setting of Gibbs measures of the form (7.1) and the relation to the Kähler-Einstein geometry will be studied in detail elsewhere [11] (for outlines see [3, 10]). Here we will just give a brief impressionistic view of the setting. The general geometric background data consists of a pair \((\mu_0, \phi_0)\) where \(\mu_0\) is a measure on the \(n\)-dimensional complex manifold \(X\) and \(\phi_0\) is a metric on an ample line bundle \(L \to X\) (more precisely, we will denote by \(\phi_0\) the collection of local functions such that \(e^{-\phi_0}\) represents the metric with respect to given local trivializations of \(L\)). To this data one may associate a sequence of Gibbs measure of the form (7.1) but with \(\Delta^{(N_k)}\) replaced with any generator of the determinant line \(\Lambda^N H^0(X, L^\otimes k)\), where \(H^0(X, L^\otimes k)\) denotes the \(N\)-dimensional space of global holomorphic sections with values in the \(k\)th tensor power of \(L\). When \(\beta_N = \beta\) the corresponding mean field type equations are then of the form

\[
MA_C(\phi) = e^{\beta(\phi - \phi_0)} \mu_0
\]

for a positively curved metric \(\phi\) on the line bundle \(L\). The relation to Kähler-Einstein geometry stems from the fact when \(L\) is taken as the canonical line bundle \(K_X := \Lambda^n (T^* X)\) any metric \(\phi_0\) determines a measure \(\mu_0 = e^{+\phi_0} dz \wedge d\bar{z}\) and the equation (7.4) is then intrinsically defined for \(\beta = 1\) (i.e. independent of \(\phi_0\)). In fact, as is well-known the equation is then equivalent to the Einstein equation with cosmological constant \(\Lambda = -1\) for the Kähler metric \(\omega := \frac{1}{2\pi} \partial \bar{\partial} \phi\) on \(X\), i.e. the equation

\[
\text{Ric } \omega = \Lambda \omega,
\]

where \(\text{Ric } \omega\) denotes the Ricci curvature of \(\omega\). Moreover, the corresponding Gibbs measure is then also intrinsically defined by \(X\). Similarly, if \(X\) is a Fano manifold, i.e. the anti-canonical line bundle, \(K^{-1}_X := \Lambda^n (TX)\) is ample, then we can take \(L = K^{-1}_X\) and for \(\beta = -1\) the corresponding equation (7.4) coincides with the Einstein equation for \(\omega\) with cosmological constant \(\Lambda = +1\). However, in this setting the corresponding Gibbs measure will not be well-defined in general since the partition function may diverge (the reason is that the corresponding integrand is then locally of the form \(1/|f_k(z_1, ..., z_N)|^{2/k}\) for a holomorphic function \(f_k\) and the integrability properties are thus reflected in the singularities of the hypersurface cut out by \(f_k\)). It is then natural to define a statistical mechanical notion of stability of a Fano manifold \(X\) called Gibbs stability by demanding that the partition function be finite for \(k\) sufficiently large [3, 11]. This should be thought of as a probabilistic version of other notions of algebro-geometric stability, such as K-stability, appearing in Kähler-Einstein geometry. Interestingly, Gibbs stability admits a purely algebro-geometric interpretation saying that the anti-canonical incidence divisor in \(X^{N_k}\) has, for \(k\) sufficiently large, mild singularities in the sense of the Minimal Model Program (or more precisely, Kawamata log terminal singularities). There are also variations of the notions of Gibbs stability, which for example, are needed when \(X\) admits non-trivial holomorphic vector fields (since \(X\) will never be Gibbs stable then). In the following sections we will outline some concrete relations between the general complex geometric setting and the present one.

7.2. A general LDP for Gibbs measures and Coulomb type gases. Let us consider the following general setting: \(X\) is a topological space equipped with a
(Borel) measure $\mu_0$ and $H^{(N)}$ is a sequence of symmetric functions on $X^N$. We also assume given a sequence $\beta_N \to \infty$ such that $Z_{N,\beta_N} := \int_{X^N} e^{-\beta_N H^{(N)} \mu_0^{\otimes N}}$ is finite for any $N$. Then the corresponding Gibbs measures $\mu^{(N)}_{\beta_N}$ are well-defined. For simplicity we will assume that $X$ is compact and that $\mu_0$ has finite mass and may hence (up to a harmless scaling) be assumed to be a probability measure. Let us also assume that the assumptions in the Gärtner-Ellis theorem hold, i.e. that there exists a Gateaux differentiable functional $F(u)$ on $C^0(X)$ such that

$$- \lim_{N \to \infty} \frac{1}{\beta N} \log Z_{N,\beta_N}[u] = F(u)$$

By the Gärtner-Ellis theorem it then follows that law of the empirical measure, i.e. $(\delta_N)_*(\mu^{(N)}_{\beta_N})$, satisfies a LDP on $\mathcal{M}_1(X)$ with speed $\beta N$ and rate functional equal to $E$ (up to an additive normalizing constant), where, as before, the functional $E(\mu)$ denotes the Legendre transform of $F$, i.e. $E = F^*$.

Replacing the sequence $\beta_N$ with a fixed positive number $\beta$ it is natural to ask under what conditions the corresponding Gibbs measures $\mu^{(N)}_{\beta}$ satisfy a LDP with rate functional equal to $E + D_{\mu_0}/\beta$ (up to an additive normalizing constant)? By Theorem 1.4 it is enough to assume equicontinuity of $H^{(N)}$, but the proof given above actually applies in a considerable more general setting and, loosely speaking, shows that the result holds as long as a certain chaoticity property holds (a more direct proof under the equicontinuity setting is given in section 4.6 below). In order to formulate this properly let us denote by $\mu^{(N)}_{\beta}$ the “tilted” Gibbs measures obtained by replacing $H^{(N)}/N$ with $H^{(N)}/N + u$. By the Gärtner-Ellis theorem $\mu^{(N)}_{\beta}$ satisfies a LDP principle for any fixed $u$ and it also follows that the $j$ th marginal of $\mu^{(N)}_{\beta}$ converges to $(\mu_u)^{\otimes j}$ where

$$\mu_u := d\mathcal{F}[u]$$

In the terminology of Kac this says that the sequence $\mu^{(N)}_{\beta}$ is $\mu_u$-chaotic (see 57 and references therein). Now the main extra property that is needed to deduce the LDP for a fixed $\beta$ is that the whole measure $\mu^{(N)}_{\beta}$ is sufficiently close to the corresponding product measure $(\mu_u)^{\otimes N}$ in an entropic sense.

**Theorem 7.1.** Assume given data $(X, \mu_0, H^{(N)})$, as above, such that the corresponding functional $\mathcal{F}(u)$ is Gateaux differentiable. Assume that

- $\inf_{\beta} \frac{1}{\beta N} \log Z_{N,\beta_N}[u] + o(1)$ for any fixed $u \in C^0(X)$
- $\lim_{N \to \infty} D_{\mathcal{F}_{\mu_0}}(\mu^{(N)}_{\beta_N}, u) = 0$ for any “good” function $u$ in $C^0(X)$ i.e. such that $D_{\mu_0}(\mu_u) < \infty$ and $E_{\mu_0}(\mu_u) < \infty$.
- There exists some probability measure $\mu$ such that $D_{\mu_0}(\mu) < \infty$ and $E(\mu) < \infty$ and any such measure may be written as a weak limit of measures $\mu_{u_j}$ with $u_j$ good such that the functionals $D_{\mu_0}$ and $E$ are continuous along $\mu_{u_j}$.

Then $\mu^{(N)}_{\beta}$ satisfies an LDP with speed $\beta N$ and rate functional $F_{\beta} := E + D_{\mu_0}/\beta$ (up to an additive constant) and

$$(7.6) \quad - \lim_{N \to \infty} \frac{1}{\beta N} \log Z_{N,\beta_N} = \inf_{\mu \in \mathcal{M}_1(X)} F_{\beta}(\mu)$$
Proof. The proof is essentially contained in the previous arguments, so we will only recall the main points. First, fixing a probability measure \( \mu \) on \( X \) and a continuous function \( u \) we may, in a similar manner as in the proof of Theorem 2.1, rewrite
\[
\int_{X^N} \frac{H^{(N)}}{N} \mu^\otimes N = \left( -\frac{1}{\beta N} \log Z_{N,\beta N} [u] - \int_X u \mu \right) + \frac{D \left( \left( \mu \right)^\otimes N, \mu^{(N)}_{\beta N} \right) - 1}{\beta N} D_{\mu_0}(\mu)
\]
(assuming that all terms are finite). In particular, if \( u \) is good and \( \mu = \mu_u \) the rhs converges to \( E(\mu) \), as \( N \to \infty \). We can then proceed exactly as before, using the Gibbs variational principle to get, for any fixed \( \mu \), by first taking the infimum over all good \( u \), that
\[
F^{(N)}_{\beta} (\mu^{(N)}_{\beta}) \leq \inf_u F_{\beta} (\mu_u) = \inf_{M_1(X)} F_{\beta} (\mu_u),
\]
where we have used the third assumption in the last equality. As for the lower bound it is proved exactly as before: if \( \mu^* \) is a limit point of the first marginal \( \mu^{(N)}_1 \), then, for any \( u \in C^0(X) \) we get
\[
\lim_{N \to \infty} \inf_X \left( \frac{H^{(N)}}{N} + u \mu^* \right) \leq \lim_{N \to \infty} F^{(N)}_{\beta} (\mu^{(N)}_{\beta})
\]
and we can then conclude that the asymptotics in formula 7.6 hold. The large deviation property then follows precisely as before by replacing \( H^{(N)} \) with \( H^{(N)} + u \).
\[\square\]

The reason that we have invoked the approximation property is that, in general, the solution of the corresponding mean field type equations may not be continuous. Anyway, in many cases continuity of the solution is guaranteed and then the approximation assumption appearing in the third point above is not needed. It should also be pointed out that the first assumption may be removed if one instead defines the functional \( F \) in terms of the infimum over \( X^N \) (as in formula 2.2). This is particularly useful when \( \mu_0 \) is very irregular. To illustrate this we state the following general result about Coulomb gases whose complete proof will appear elsewhere.

**Theorem 7.2.** Let \( \mu_0 \) be a (Borel) measure on \( C \) and \( \phi_0 \) a continuous function of super logarithmic growth (i.e. \( \phi_0 \geq \log((1 + |z|^2) - C) \) defined on the support of \( \mu_0 \), such that \( \int e^{-\beta \phi_0} \mu_0 < \infty \). Set
\[
H^{(N)}(z_1, \ldots, z_N) := -\frac{1}{N} \sum_{1 \leq i \neq j \leq N} \log |z_i - z_j| + \sum \phi(z_i)
\]
and
\[
E(\mu) := -\int \log |x - y| \mu(x) \otimes \mu(y) + \int \phi_0 \mu
\]
Then, for any positive number \( \beta \), the law of the empirical measure of the corresponding Gibbs measure \( \mu^{(N)}_{\beta} \) satisfies a LDP with rate functional \( E + D_{\mu_0}(\mu)/\beta \) (up to an additive constant) as long as the latter functional is not identically equal to infinity.

We recall that corresponding Gibbs measure \( \mu^{(N)}_{\beta} \) may in this setting be written in terms of the Vandermonde determinant for polynomials of degree \( \leq k := N - 1 \)
(compare section 7.1 for the relation to permanents):
\[
\mu_{\beta}^{(N)} = \frac{1}{Z_{N,\beta}} |\Delta^{(N)}(z_1, \ldots, z_N)|^{2\beta/k} e^{-\beta(\phi(z_1) + \cdots + \phi(z_N))} \mu_0 \otimes \mu_0
\]

(this is a special case of the complex geometric framework outlined i the previous section, where \( X \) is the complex projective line (i.e. the Riemann sphere) \( \mathbb{P}^1 \), viewed as the one-point compactification of \( \mathbb{C} \), and \( L \) is the hyperplane line bundle \( \mathcal{O}(1) \to \mathbb{P}^1 \)). In this setting the corresponding mean field equations on \( \mathbb{C} \) may, in complex notation, be written as
\[
\frac{i}{2\pi} \partial \bar{\partial} \phi = e^{\beta(\phi - \phi_0)} \mu_0,
\]
where the function \( \phi \) (which is automatically subharmonic) satisfies the normalization condition \( \frac{i}{2\pi} \int_\mathbb{C} \partial \bar{\partial} \phi = 1 \) (which for example holds if \( \phi = \log |z|^2 + O(1) \) as \( |z| \to \infty \)).

In the case when \( \mu_0 \) is the Lebesgue measure supported on a bounded domain \( \Omega \) in the plane, the previous theorem was first shown in [18, 30]. But the main point here is that the method of proof indicated above applies to any measure \( \mu_0 \) with the property that the corresponding free energy functional \( F_\beta \) is not identically equal to infinity (which is equivalent to the existence of a minimizer). For example, any measure \( \mu_0 \) not charging polar sets will do. In particular, the assumption about the Bernstein-Markov property of \( \mu_0 \) which appears in the case \( \beta = \infty \) [6, 13] is not needed for \( \beta \) finite.

However, in the case of \( \beta < 0 \) stronger assumptions on \( \mu_0 \) are needed. It turns out that an essentially optimal regularity assumption is that there are positive constants \( C \) and \( d \) such that the measure \( \mu_0 \) satisfies
\[
\mu_0(B_r) \leq Cr^d,
\]
for \( r \) sufficiently small, for every Euclidean ball of radius \( r \). Under this assumption the corresponding Gibbs measure \( \mu_{\beta}^{(N)} \) is well-defined for any \( \beta > -d \) and, after passing to a subsequence, the law of the corresponding empirical measure converges weakly to a measure concentrated on the set of minimizers of \( F_\beta \) [11]. In particular situations the uniqueness of such minimizers can be ensured. See [18, 30] for the case when \( \mu_0 \) is the Lebesgue measure supported on a bounded domain \( \Omega \) in the plane. The authors main motivation for studying this setting comes from the relation to the study of Kähler-Einstein metrics on complex algebraic varieties, i.e. Kähler metrics with constant Ricci curvature. For example, the general case referred to above in particular applies to the following setting which corresponds to the complex geometric setting of of conical metrics on the Riemann sphere with positive constant curvature:

**Theorem 7.3.** [11] Consider \( m \) points \( p_1, \ldots, p_m \) in \( \mathbb{C} \) with weights \( w_1, \ldots, w_m \in [0, 1] \) such that \( d := 2 - (w_1 + \cdots + w_m) > 0 \) and set \( \phi_0 = 0 \) and
\[
\mu_0 := \left| \frac{1}{z_1 - p_1} \cdots \frac{1}{z_m - p_m} \right|^{2w_1} \cdots \left( \frac{1}{2} \right) \frac{dz \wedge d\bar{z}}{2w_m}
\]
Then the following is equivalent:
- The equation (7.7) admits a unique solution for \( \beta = -d \)
- the corresponding partition functions \( Z_{N,-d} \) are finite for \( N \gg 1 \)
Moreover, in that case the corresponding empirical measure converges in probability to $\mu_{-d} := \frac{1}{2\pi} \partial \bar{\partial} \phi_{-d}$ where $\phi_{-d}$ is the solution of equation (7.7).

The relation to complex geometry comes from the well-known fact that $\phi$ is a solution to the equation appearing in the previous theorem iff $\omega := \frac{1}{2\pi} \partial \bar{\partial} \phi$ defines a Kähler metric on the Riemann sphere $X$ with constant positive curvature and conical singularities encoded by the effective $\mathbb{R}$-divisor $E := p_1 w_1 + \cdots + p_m w_m$, i.e. $\omega$ is a conical Kähler-Einstein metric. From this point of view the previous theorem can be formulated as saying that a one dimensional log Fano variety $(X, E)$ admits a unique conical Kähler-Einstein metric iff it is Gibbs stable in the sense of [10]. This and the higher dimensional setting will be studied in detail in [11]. For the moment we just point out that it is well-known that a necessary condition for uniqueness in the previous theorem is that there are at most two points $p_i$ or more precisely, either (i) no points or (ii) precisely two points (and in the latter case $w_1 = w_2$). In the first case uniqueness indeed fails since the equations are invariant under all biholomoprhic maps of the Riemann sphere (i.e. the Möbius group) and in the second case invariance holds under the standard $\mathbb{C}^*$ action which fixes the two points. This symmetry can also be seen to be responsible for the fact that the corresponding partition functions $Z_N$ then diverge. However, there is a way to break the symmetry in order to restore uniqueness and finiteness in the previous theorem. To this end one fixes a subharmonic function $\phi$ that there is a unique solution $\phi$ to the corresponding mean field type equations for $\beta > -d$. Moreover, the corresponding partition functions $Z_{N,-\beta}$ are then finite and the empirical measure converges in probability towards corresponding measure $\frac{1}{2\pi} \partial \bar{\partial} \phi_{-d}$. In the particular case when $\phi_0$ has circular symmetry this is closely related to the one-dimensional real Monge-Ampère equations. In fact, a similar phenomena persists in higher dimensions under toric symmetry. This is the subject of the the next section where we will outline the relation between toric Kähler-Einstein metrics and the previous probabilistic setting of permanents and the real Monge-Ampère equation.

7.3. Toric Kähler-Einstein metrics, negative $\beta$ and phase transitions. First consider the following special case of the setting of weighted measures $(\mu_0, \phi_0)$ in $\mathbb{R}^n$ (section 5.11): given a weight function $\phi_0$ on $\mathbb{R}^n$ we take the measure $\mu_0$ to be given by $\mu_{\phi_0} := e^{-\phi_0} dx$. Setting $\gamma = -\beta$ the corresponding Monge-Ampère equation (7.8) may then be written as

$$MA(\phi) = e^{-(\gamma + (1-\gamma)\phi_0)} dx,$$

which in turn can be written as a twisted Kähler-Einstein equation on the complex torus $\mathbb{C}^n$. Indeed, let $\text{Log}$ be the map from $\mathbb{C}^n$ to $\mathbb{R}^n$ defined in section 7.1 and set $\varphi := \text{Log}^* \phi$. Then $\omega := \frac{1}{2\pi} \partial \bar{\partial} \varphi$ defines a Kähler metric on $\mathbb{C}^n$ which satisfies

$$\text{Ric } \omega = \gamma \omega + (1 - \gamma)\omega_0,$$

where $\text{Ric } \omega$ is the Ricci curvature of the Kähler metric $\omega$, represented as two-form. In particular, for $\gamma = 1$ a solution $\omega$ is a bona fide Kähler-Einstein metric, i.e. a Kähler metric with constant (positive) Ricci curvature and the corresponding convex function $\phi(x)$ then satisfies the $\phi_0$—independent equation

$$MA(\phi) = e^{-\phi} dx.$$
When the convex body $P$ is a polytope (containing zero in its interior) $\omega$ extends to a (singular) Kähler-Einstein metric on the corresponding toric variety $X_P$ compactifying $\mathbb{C}^n$ (see [31] and references therein). The most studied case is when $P$ is a reflexive Delzant polytope, which equivalently means that $X_P$ is a Fano manifold. Then the equation (7.11) coincides with Aubin’s continuity equation, designed by Aubin to prove the existence of a Kähler-Einstein metric by deforming $\gamma$ from $\gamma = 0$ to $\gamma = 1$. The existence of solutions for $\gamma$ a sufficiently small positive number was shown by Aubin. However, as is well-known there are in general obstructions to the existence of Kähler-Einstein metric with positive Ricci curvature and according to the fundamental Yau-Tian-Donaldson conjecture the existence of a Kähler-Einstein metrics is equivalent to an algebro-geometric notion of stability, called K-stability.

Here we will only briefly explain the relation between Aubin’s continuity equation and the probabilistic framework as developed in previous sections. To this end we assume that the given function $\phi_0$ is in $P_+^{\mathbb{R}^n}$. As shown in [31] (generalizing the seminal result of Wang-Zhu concerning the smooth Fano case) the equation (7.8) then admits a solution iff $0$ is the barycenter in $P$. More precisely, denoting by $R$ the sup over all $\gamma \in [0,1]$ such that the equation admits a solution in $P_+^{\mathbb{R}^n}$, it was shown in [9] that $R$ is given by the following formula:

$$R := \frac{\|q\|}{\|q - b\|},$$

where $q$ is the point in $\partial P$ where the line segment starting at $b$ and passing through $0$ meets $\partial P$ (this is a generalization of a result of Li concerning the smooth Fano case). Moreover, the corresponding solution is unique for $\gamma < 1$. Interestingly, the free energy functional $F_\beta$ may be identified with Mabuchi’s K-energy functional in this setting (or rather its twisted version, compare [31, 34]). To see the connection to the random point processes considered in the previous section we note that the processes (i.e. the Gibbs measures) in question are still defined for negative $\beta$ (i.e. positive $\gamma$) as long as the corresponding partition function, which may be identified with Mabuchi’s $K$-energy functional $F_\beta$ does not hold any more, but this is only a minor technical point. More seriously, since $\beta < 0$ the argument for the lower bound in the proof of Theorem 4.3 is not valid anymore as it stands. However, this problem can be circumvented using the Hewitt-Sanders decomposition theorem and the sub-additivity of the entropy, as in [34] (it is also important to know that $\mu_\beta$ is still the unique minimizer of the free energy functional $F_\beta$, which is indeed the case [9]). Another useful fact is that the first correlation measures are of the form $e^{-\gamma \phi_k + (1-\gamma)\phi_k(x)}dx$, where $\phi_k(x)$ is in $P_+^{\mathbb{R}^n}$, as follows immediately from the Prekopa inequality. This and further
relations to the Kähler-Einstein problem are deferred to [11]. Here we will only summarize the corresponding main results:

**Theorem 7.4.** [11] Let $P$ be a convex body containing 0 in its interior and $\phi_0$ a convex function on $\mathbb{R}^n$ such that $\phi_0 - \phi_P$ is bounded. Then, for any $\gamma < R$ the corresponding Gibbs measure $\mu_\gamma$ is well-defined and the law of its empirical measure converges in probability, when $N \to \infty$, towards $\mu_\gamma := MA(\phi_\gamma)$, where $\phi_\gamma$ is a solution to the equation **7.10**.

One subtle feature of this setting is the presence of translational symmetry at the critical value $\gamma = 1$ (assuming that the barycenter of $P$ vanishes) and the way that it is broken by introducing a weight $\phi_0$. The point is that for $\gamma < 1$ the corresponding mean field type equations **7.8** have a unique solution, while for $\gamma = 1$, there is an $n-$dimensional space of solutions. This is due to the fact that, in the latter case, the equations are invariant under the action of $\mathbb{R}^n$ by translations. On the other hand, introducing a weight $\phi_0$ brakes this symmetry and it turns out that the corresponding solutions $\phi_\gamma$ tend, when $\gamma \to 1$, to a particular solution $\phi$ of the equation **7.10** depending on the choice of $\phi_0$. The most transparent case is when $P$ as well as $\phi_0$ are symmetric around the origin, i.e. $-P = P$ and $\phi_0(-x) = \phi_0(x)$. Then we have the following

**Corollary 7.5.** Suppose that the convex body $P$ is symmetric and that $\phi_0$ is also symmetric with respect to the origin, i.e. $\phi_0(-x) = \phi_0(x)$. Then,

$$\phi_\gamma(N)(x) := -\frac{1}{\gamma} \log \int_{(\mathbb{R}^n)^N} \frac{1}{Z_{N,-\gamma}} (\text{Per}(x,x_2,...,x_N))^{-\gamma/k} (e^{-(1-\gamma)\phi_0} dx)^\otimes N-1,$$

converges point-wise, in the double limit where first $N \to \infty$ and then $\gamma \to 1$, to the unique solution $\phi$ of the equation **7.10** satisfying $\phi(-x) = \phi(x)$.

In the general case, it can be shown that $\phi_\gamma$ converges to the unique solution whose Monge-Ampère measure minimizes the associated energy functional $E_\phi$ on the solution space. Interestingly, the $\mathbb{R}^n-$symmetry is also responsible for the fact that at the critical value $\gamma = 1$ the corresponding random point processes are not well-defined (for any $N$). Indeed, $\text{Per}(x_1,x_2,...,x_N)$ is invariant under the diagonal action of $\mathbb{R}^n$ and hence the corresponding partition function $Z_{N,-1}$ diverges.

Let us finally point out that, in general, the critical value $\gamma = R$ can be interpreted as a second order phase transition (compare the discussion in the end of [30], which turns out to be related to the simplest case of the present setting, namely when $n = 1$ and $P = [-1,1]$).

**7.4. Langevin dynamics.** In this section we will briefly comment on a dynamical version of Theorem **7.1** which, from the point of view of statistical mechanics corresponds to the relaxation to equilibrium of the corresponding system (often referred to as Langevin dynamics). It can be seen as a fully non-linear version (when $n > 1$) of McKean’s interacting diffusions [32] [33] which concern the case when the Hamiltonian is a sum of two-point functions.

For simplicity we consider the weighted setting $(\mu_0,\phi_0)$ when $\mu_0 = dx$, so that $X = \mathbb{R}^n$ and $\phi_0$ is thus a weight function on $\mathbb{R}^n$ and we first assume that $\beta > 0$. Then we introduce the following system of Stochastic Differential Equations (SDE) for $x_1(t),...,x_N(t)$ viewed as stochastic processes with values in $\mathbb{R}^n$:
\[ dx_i(t) = -\nabla_{x_i} H_{\phi_0}^{(N)}(x_1,\ldots,x_i,\ldots x_N)dt + \frac{2}{\beta^{1/2}} dB_i(t), \]

where \( H_{\phi_0}^{(N)} \) is defined by formula (5.1) and \( \nabla_{x_i} \) denotes the (partial) gradient defined with respect to the Euclidean metric on \( \mathbb{R}^n \) and the \( B_i \)s are \( N \) independent standard Brownian motions on \( \mathbb{R}^n \). This is thus a system of Itô diffusions which can be seen as the down-ward stochastic gradient flow on \( X^N \) for the \( N \)-particle Hamiltonian \( H_{\phi_0}^{(N)} \). In concrete terms the system is obtained by adding noise to the following system of ODEs:

\[
\frac{\partial x_i}{\partial t} = \frac{1}{k} \sum_{\sigma \in S_N} \frac{\sum_{\sigma(i)} e^{x_1 P_{\sigma(1)} + x_2 P_{\sigma(2)} + \cdots + x_N P_{\sigma(N)}}}{\sum_{\sigma \in S_N} e^{x_1 P_{\sigma(1)} + x_2 P_{\sigma(2)} + \cdots + x_N P_{\sigma(N)}}} - \nabla_{x_i} \phi_0(x_i).
\]

It seems natural to conjecture that, given the initial condition that \( x_i(0) \) be i.i.d variables with law \( \rho_0 \) for some fixed, say smooth and strictly positive probability density \( \rho_0 \) on \( \mathbb{R}^n \), the system \((7.12)\) converges, when \( N \to \infty \) (in a sense to be detailed below), to a solution \( \rho_t \) of the following deterministic fully non-linear parabolic system of PDEs with initial data \( \rho_{t=0} = \rho_0 \):

\[
\frac{\partial \rho_t}{\partial t} = \frac{1}{\beta} \Delta \rho - \nabla \cdot (\rho \nabla (\phi_t - \phi_0))
\]

\[ \rho_0 dx = MA(\phi_t), \]

i.e. \( \phi_t \) is the unique normalized potential of \( \rho_t dx \) in the class \( \mathcal{P}(\mathbb{R}^n) \), which we recall means that \( \phi_t \) convex and its subgradient image is contained in the given convex body \( P \) (interestingly, a closely related parabolic system appears in dynamical meteorology - see [31] and references therein). More precisely, the converge referred to above should hold in the following sense: for a fixed time \( t \), the empirical measures \( \delta_N(x(t)) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)} \) determined by the the SDEs \((7.12)\) converge in probability to the deterministic measure \( \mu_t = \rho_t dx \). In fact, if one assumes that the empirical measures \( \delta_N(x(t)) \) converge in probability to some deterministic measure \( \mu_t \), then it can be shown, using Theorem [22], and the linear Fokker-Planck equations associated to \((7.12)\) (i.e. the corresponding forward Kolmogorov equations), that the density \( \rho_t \) of \( \mu_t \) evolves according to the parabolic PDE \((7.13)\) As is well-known, the general problem of establishing a priori convergence in probability is essentially equivalent to establishing propagation of chaos in the sense of [37] and we leave this problem for the future.

In the light of the discussion in the previous section and the connections to Kähler-Einstein metrics on toric varieties it is also very interesting to study the case when \( \beta < 0 \) where one would expect that there exists a global solution to the system \((7.13)\) in the case when \( \beta \geq -R \), where \( R \) is the invariant of the convex body \( P \) defined by formula \((7.11)\) and that the convergence statement should hold for \( \beta > -R \). Finally, it may be worth pointing out that, inspired by ideas introduced by Otto (see [38] and references therein), it can be shown that the parabolic equation \((7.13)\) is (at least formally) the down-ward gradient flow for the corresponding free energy functional \( F_\beta \), defined with respect to the Wasserstein 2-metric on the space \( \mathcal{M}_1(\mathbb{R}^n) \), when \( \mathbb{R}^n \) is equipped with the Euclidean metric. This observation becomes particularly striking in the toric setting considered in the previous section,
where $F_\beta$ may be identified with the Mabuchi K-energy functional. In fact, for any Kähler manifold $(X, \omega)$ the down-ward gradient flow of the latter functional, defined with another metric, namely the one defined by the Mabuchi-Semmes-Donaldson metric on the space of all Kähler metrics in the Kähler class $[\omega]$ is precisely the Calabi flow which plays a prominent role in Kähler geometry. This also motivates studying the complex version of the parabolic equation 7.13 which is naturally defined on any given Kähler manifold $(X, \omega)$. In fact, the complex version of 7.13 (for $\beta$ negative) in the case when $X$ is Riemann sphere is closely related to the Keller-Segal system in $\mathbb{R}^2$, which has been extensively studied in recent years (see for example [12] and references therein). It seems also natural to conjecture that the complex version of the parabolic equation 7.13 may be obtained as the large $N$-limit of a systems of SDE’s of the form 7.12 obtained by replacing the permanent appearing in the definition of $H^{(N)}_{\phi_0}$ by the corresponding Vandermonde type determinant (compare [6, 10, 3, 8]). But we also leave the study of this complex story for the future.

8. Appendix

A1: Background on optimal transport and its discrete version. The classical assignment problem (also known as the bivariate perfect matching problem in graph theory) is the problem to, given an $N \times N$ matrix $(c_{ij})$ minimize the functional

$$\sigma \mapsto \sum_{i=1}^{N} c_{i\sigma(i)}$$

In economical terms we have $N$ workers and $N$ jobs to conduct and $c_{ij}$ is the cost of assigning work $j$ to a worker $i$. The problem is to minimize the total cost, if all the every workers are assigned different jobs, i.e. worker $i$ is assigned the job $j$ where $j = \sigma(i)$ for some permutation $\sigma \in S_N$.

The assignment problem relevant to the present paper appears in the following setting of (discrete) optimal transport theory. Consider two sets $X$ and $P$ in $\mathbb{R}^n$ and a given cost function $c(x, p)$ on $X \times P$. As in the previous sections we denote by $N$ the number of lattice points in $P$, i.e. the points in $P_\mathbb{Z} := P \cap \mathbb{Z}^n$. Fix also a configuration $(x_1, \ldots, x_N)$ of $N$ points on $X$. Then we define the transport cost from $x_i$ to $p_j$ as the number $c_{ij} := c(x_i, p_j)$. Fixing an order $p_1, \ldots, p_N$ of the points in $P_\mathbb{Z}$ the problem of (discrete) optimal transport may then be defined as the corresponding assignment problem, i.e. as the problem of minimizing (8.1) for $c_{ij} := c(x_i, p_j)$. Concretely, this means that we want to assign $N$ different points in $P$ to the $N$ given points $x_i$ on $X$ in such a way that the corresponding total cost is minimized. As before we get an asymptotic problem, with $N$ tending to infinity, by replacing $P_\mathbb{Z}$ with $P_{\mathbb{Z}/k} := P \cap (\mathbb{Z}/k)^n$ to get a sequence of discrete optimal transport problems. When studying asymptotics it will be convenient to divide the total cost (appearing formula (8.1)), by the number $N$ of workers to get the average cost of the work performed. Accordingly, we define the (normalized) cost $C(\sigma)$ by

$$C(\sigma) := \frac{1}{N} \sum_{i=1}^{N} c(x_i, p_{\sigma(i)})$$
8.0.1. Optimal transport theory (continuous version). In the classical “continuous” setting for optimal transport theory, as originally introduced by Monge, the given data consist of two probability measures \( \mu \) and \( \nu \) on \( \mathbb{R}^n \) and a cost function \( c(x, p) \) on \( \mathbb{R}^n \times \mathbb{R}^n \) (see the monographs [38, 39] for further background and extensive references). A transport map \( T \) is, by definition, a map from \( \mathbb{R}^n \) to \( \mathbb{R}^n \) such that
\[
T_\ast \mu = \nu
\]
One defines the transport cost of the transport map \( T \) as
\[
c(T) := \int_{\mathbb{R}^n} c(x, T(x)) \mu
\]
and \( T \) is said to be an optimal transport map if it minimizes the cost \( c(T) \) over all transport maps (i.e. those satisfying the push-forward formula \( (8.3) \)). However, in general such an optimal transport map may not exist and following Kantorovich one usually considers a relaxed version of Monge’s problem where the transport map \( T \) is replaced with a coupling \( \Gamma \) (between \( \mu \) and \( \nu \)) i.e. \( \Gamma \) is a measure on \( \mathbb{R}^n \times \mathbb{R}^n \) whose push-forwards to the first and second factor are equal to \( \mu \) and \( \nu \), respectively (such a \( \Gamma \) is also called a transference plan). Its cost is then defined by
\[
C(\Gamma) := \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, p) \Gamma
\]
which is thus the restriction of a linear functional to the space of all couplings (an optimal coupling \( \Gamma \) exists under very general assumptions [38, 39]). Accordingly, fixing \( \nu \), the optimal total cost to transport \( \mu \) to \( \nu \) is defined by
\[
C(\mu) := C(\mu, \nu) := \inf_{\Gamma} c(\Gamma),
\]
where the infimum is taken over all couplings between \( \mu \) and \( \nu \). In particular, any transport map \( T \) defines a coupling \( \Gamma_T := (I \times T)_\ast \mu \) such that \( C(T) = C(\Gamma_T) \).

Assume now that we are given a closed set \( X \) in \( \mathbb{R}^n \) and a convex body \( P \) of unit-mass. We then fix \( \nu := \lambda_P \) to be the Lebesgue measure supported on \( P \) (that we will sometimes also write as \( 1_P dp \)) and consider cost functions \( c(x, p) \) of the form
\[
(8.4) \quad c_{\phi_0}(x, p) := -x \cdot p + \phi_0(x)
\]
where \( \phi_0 \) is a given weight function on \( X \) (compare section 5.1.1) and denote by \( C_{\phi_0}(\mu) \) the corresponding optimal cost functional. It may be decomposed as
\[
C_{\phi_0}(\mu) = C_0(\mu) + \int \phi_0 \mu,
\]
where \( C_0 \) is the “unweighted” cost functional defined with respect to \( c(x, p) := -x \cdot p \). Since we are only interested in the dependence of \( C_{\phi_0}(\mu) \) with respect to \( \mu \) we could also have added any continuous function \( \psi_0(p) \) to the cost function \( c_{\phi_0}(x, p) \). Indeed, this would only shift \( C_{\phi_0}(\mu) \) by an overall additive constant. A classical case is when \( \phi_0(t) = \psi_0(t) = |t|^2/2 \), so that the correspondind cost function is \( |x - p|^2 \) and \( C(\mu) \) is hence the Wasserstein 2-distance between \( \mu \) and \( \lambda_P \).

To see the relation to the discrete setting above we note that for any given cost function \( c(x, p) \) setting \( \mu := \delta(x^{(N)}) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \) for a given configuration \((x_1, ..., x_N)\) of points on \( X \) and \( \nu := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \) clearly gives,
\[
(8.5) \quad C(\Gamma) = C(\sigma)
\]
Given a ball $B_1$ the first equality in the lemma can be proved by various means (for example
Proof. inequality only using the fact that $v$ holds for any $\tau$ by Stokes theorem and a simple approximation argument the latter integral is equal more generally elements in $\epsilon > 0$ constant (Comparison principle) Let $\mu$ in $P_1(X)$, the previous equality says that the embedding [2.34] is an isometry.

**A2: The comparison and domination principles for MA.** In this appendix we will provide proofs of the comparison and domination principle for the Monge-Ampère operator acting on the function space $P(\mathbb{R}^n)$ associated to a convex body $P$ (following the notation in section [5.1]). These results are without doubt well-known to experts, but for completeness we have provided proofs that mimic the proofs in the complex setting (see [13] and references therein). In fact, the proofs only use the following basic properties of the real Monge-Ampère operator $MA$:

- $MA$ is a local operator on $P(\mathbb{R}^n)$, i.e. if $u = v$ on an open set $U$ then $1_U MA(u) = 1_U MA(v)$.
- For any $u \in P_+(\mathbb{R}^n)$ the measure $MA(u)$ is a probability measure.
- The space $P_+(\mathbb{R}^n)$ is closed under the max operation

One reason for isolating these ambient properties is that they may be useful when studying the general space $P(Y)$ of “ambient potentials” associated to a Hamiltonian as in section [4.5] but we will not go further into this here. We start with a verification of the second part of the first point above:

**Lemma 8.1.** Let $u$ and $v$ be elements in $P(\mathbb{R}^n)$ of maximal growth, i.e. $u, v$ are in $P_+(\mathbb{R}^n)$, which by definition means that $u - \phi_P$ and $v - \phi_P$ are bounded. Then

$$\int_{\mathbb{R}^n} MA(v) = \int_{\mathbb{R}^n} MA(u) = 1.$$

More generally, if $u$ and $v$ be elements in $P(\mathbb{R}^n)$ such that $u \to \infty$ as $|x| \to \infty$ and $v \leq u + C$ on $\mathbb{R}^n$ for some constant $C$, then

$$\int_{\mathbb{R}^n} MA(v) \leq \int_{\mathbb{R}^n} MA(u).$$

**Proof.** The first equality in the lemma can be proved by various means (for example using the Legendre transform). Here we will instead prove the more general second inequality only using the fact that $MA$ satisfies a weak form of Stokes theorem.

Given a ball $B_{R_1}$ of radius $R_1$ centered at 0 there exists a constant $R_2 > R_1$ and constant $\epsilon > 0$ and $A > 0$ such that $(1 - \epsilon)v + A \geq 0$ on $B_{R_1}$ and $(1 - \epsilon)v + A \leq u$ on $B_R$ for any $R > R_2$. Hence, setting $\tilde{v} := \max\{(1 - \epsilon)v + A, u\}$ gives $\tilde{v} = v$ on $B_{R_1}$ and $u$ on $B_R$ for $R > R_1$. Hence, $\int_{B_{R_1}} MA((1 - \epsilon)v) = \int_{B_{R_1}} MA(\tilde{v}) \leq \int_{B_R} MA(\tilde{v})$. But, by Stokes theorem and a simple approximation argument the latter integral is equal to $\int_{B_R} MA(u)$ and hence $\int_{B_{R_1}} MA(v(1 - \epsilon)) \leq \int_{B_R} MA(u)$. Since, this inequality holds for any $R_1 > 0$ and $\epsilon > 0$ this concludes the proof of the lemma.

**Proposition 8.2.** (Comparison principle) Let $u$ and $v$ be elements in $P_+(\mathbb{R}^n)$ (or more generally elements in $P(\mathbb{R}^n)$ of full Monge-Ampère mass). Then

$$\int_{\{u < v\}} MA(v) \leq \int_{\{u < v\}} MA(u).$$
Proof. Let us first prove the a priori weaker inequality

\[
\int_{\{u<v\}} MA(v) \leq \int_{\{u\leq v\}} MA(u).
\]

Since \( MA \) is a local operator we have

\[
1_{\{u<v\}} MA(v) = 1_{\{u<v\}} MA(\max(u,v)) \leq 1_{\{u\leq v\}} MA(\max(u,v)).
\]

Writing \( \{u \leq v\} = \mathbb{R}^n - \{u > v\} \) and using locality again hence gives

\[
1_{\{u<v\}} MA(v) \leq I_{\mathbb{R}^n} MA(\max(u,v)) - 1_{\{u<v\}} MA(u)
\]

Integrating this inequality over \( \mathbb{R}^n \) and using that, by assumption, \( \int_{\mathbb{R}^n} MA(u) = \int_{\mathbb{R}^n} MA(v) \) then gives

\[
\int_{\{u<v\}} MA(v) \leq \int_{\mathbb{R}^n} MA(u) - \int_{\{u<v\}} MA(u)
\]

which hence proves the inequality \( \ref{inequality} \). Finally, to treat the general case we apply the previous inequality to \( u+\delta \) and \( v \) giving

\[
\int_{\{u+\delta<v\}} MA(v) \leq \int_{\{u+\delta\leq v\}} MA(u).
\]

Finally, letting \( \delta \to 0 \) and using that the two sequences of sets \( \{u+\delta<v\} \) and \( \{u+\delta \leq v\} \) both increase to \( \{u<v\} \) concludes the proof of the proposition. \( \square \)

Now we can prove the following

**Corollary 8.3.** (Domination principle) Let \( u \) and \( v \) be elements in \( P(\mathbb{R}^n) \) such that \( u \) is in \( P_+ (\mathbb{R}^n) \). If \( u \geq v \) almost everywhere with respect to the Monge-Ampère measure \( MA(u) \) then \( u \geq v \) everywhere on \( \mathbb{R}^n \).

**Proof.** First note that we may as well assume that \( v \) is also in \( P_+(\mathbb{R}^n) \), by replacing \( u \) with \( \max\{u,v\} \). In the case when \( MA(v) > \delta dx \) for some \( \delta > 0 \) the corollary follows immediately from the previous proposition. In the general case we simply fix an element \( v_+ \) in \( P(\mathbb{R}^n) \) such that \( MA(v_+) > \delta dx \) for some \( \delta \) and \( v_+ \leq u \) on \( \mathbb{R}^n \) (for example, \( v_+ = \log \int_{\mathbb{R}^n} e^{v+} dp - C \) for \( C \) sufficiently large) and apply the previous argument to \( u \) and \( v_+ := (1-\epsilon)u + \epsilon v_+ \) for any \( \epsilon > 0 \). This shows that \( v_+ \leq u \) on \( \mathbb{R}^n \) and letting \( \epsilon \to 0 \) thus concludes the proof. \( \square \)

**References**

[1] Bakelman, I.J: Convex analysis and non-linear geometric elliptic equations. Springer-Verlag (1994)

[2] Berman, R.J: Bergman kernels for weighted polynomials and weighted equilibrium measures of \( C^-n \). 19 pages. Indiana University Mathematics Journal, Volume 58, issue 4, 2009

[3] Berman, R.J; Kahler-Einstein metrics emerging from free fermions and statistical mechanics. J. of High Energy Physics (2011).

[4] Berman, R.J; Boucksom, S: Growth of balls of holomorphic sections and energy at equilibrium. Invent. Math. 181 (2010), no. 2, 337–394.

[5] Berman, R.J.; Boucksom, S; Witt Nyström, D: Fekete points and equidistribution on complex manifolds. Acta Math. Vol. 207, Issue 1 (2011), 1-27

[6] Berman, R.J; Determinantal point processes and fermions on complex manifolds: Large deviations and Bosonization. [arXiv:0812.3224]

[7] Berman, R.J.; Boucksom, S; Guedj, V; Zeriahi, A: A variational approach to complex Monge-Ampère equations. Publications mathématiques de l’IHÉS (to appear). [arXiv:0907.4490]

[8] Berman, R.J: A thermodynamical formalism for Monge-Ampere equations, Moser-Trudinger inequalities and Kahler-Einstein metrics. [arXiv:1011.3076]
[9] Berman, R.J; Berndtsson, B: Real Monge-Ampere equations and Kahler-Ricci solitons on toric log Fano varieties. arXiv:1209.0996
[10] Berman, R.J: A probabilistic approach to Kahler-Einstein metrics. http://sms.cam.ac.uk/media/1245708. Video and audio from a workshop on Kahler Geometry in Cambridge 2012, organized by I.Cheltsov and J.Ross
[11] Berman, R.J: A probabilistic approach to Kähler-Einstein metrics, stability and Coulomb type gases. Article in preparation.
[12] Blanchet, A; Carlen, Eric A.; Carrillo, J. A. Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. J. Funct. Anal. 262 (2012), no. 5, 2142–2230.
[13] Bloom, T; Levenberg, N. Pluripotential energy and large deviation. arXiv:1110.6593
[14] Boucksom, S; Eyssidieux, P; Guedj, V; Zeriahi, A: Monge-Ampère equations in big cohomology classes. Acta Math. 205 (2010), no. 2, 199–262.
[15] Bovier, A: Statistical Mechanics of Disordered Systems: A Mathematical Perspective. Cambridge University Press.
[16] Boucher, C; Ellis, R. S.; Turkington, B: Derivation of maximum entropy principles in two-dimensional turbulence via large deviations. (English summary) J. Statist. Phys. 98 (2000), no. 5-6,
[17] Brenier, Y: Polar factorization and monotone rearrangement of vector-valued functions. Comm. Pure Appl. Math. 44 (1991), no. 4, 375–417.
[18] Caglioti,E; Lions, P-L; Marchioro.C; Pulvirenti.M: A special class of stationary flows for two-dimensional Euler equations: a statistical mechanics description. Communications in Mathematical Physics (1992) Volume 143, Number 3, 501-525
[19] Donaldson, S. K. Some numerical results in complex differential geometry. Pure Appl. Math. Q. 5 (2009), no. 2.
[20] Caffarelli, L.A: Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. Ann. of Math. (2) 131 (1990), no. 1, 135–150.
[21] Caffarelli, L.A: Some regularity properties of solutions of Monge Ampère equation. Comm. Pure Appl. Math. 44 (1991), no. 8-9, 965–969.
[22] Caffarelli, L. A. A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. Ann. of Math. (2) 131 (1990), no. 1, 129–134.
[23] Caffarelli, L.A: The regularity of mappings with a convex potential. J. Amer. Math. Soc. 5 (1992), no. 1, 99–104.
[24] Dembo, A; Zeitouni, O: Large deviations techniques and applications. Jones and Bartlett Publishers, Boston, MA, 1993. xiv+346 pp.
[25] Ellis, R. S.; Haven, K; Turkington, B: Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. J. Statist. Phys. 101 (2000), no. 5-6, 999–1064.
[26] Gao, W; McCann, R. J.: The geometry of optimal transportation. Acta Math. 177 (1996), no. 2, 113–161.
[27] Gutierrez, C.E: The Monge-Ampère equation. Progress in Nonlinear Differential Equations and their Applications, 44. Birkhäuser Boston, Inc., Boston, MA, 2001. xii+127 pp. ISBN: 0-8176-4177-7
[28] Hough, J. B.; Krishnapur, M.; Peres, Y.l; Virág, B: Determinantal processes and independence. Probab. Surv. 3 (2006), 206–229
[29] Huesmann, M; Sturm, K-T: Optimal Transport from Lebesgue to Poisson. http://arxiv.org/abs/1012.3845
[30] Kiesling M.K.H.: Statistical mechanics of classical particles with logarithmic interactions, Comm. Pure Appl. Math. 46 (1993), 27-56.
[31] Loeper, G: A fully non-linear version of Euler incompressible equations: the Semi-Geostrophic system. SIAM Journal of Math Analysis (to appear).
[32] McKean, H. P. Jr. A class of Markov processes associated with nonlinear parabolic equations. Proceedings of the National Academy of Science 56: 1907-1911, 1966.
[33] McKean, H. P. Jr. Propagation of chaos for a class of nonlinear parabolic equations. Lecture Series in Di®erential Equations 7: 41-57. Catholic University, Washington, D.C., 1967.
[34] Messer, J; Spohn, H: Statistical mechanics of the isothermal Lane-Emden equation. J. Statist. Phys. 29 (1982), no. 3, 561–578.
[35] Negele, J.W; Orland, H: Quantum Many Particle Systems. Westview Press (1998)
[36] Rockafellar, R. T: Convex analysis. Reprint of the 1970 original. Princeton Landmarks in Mathematics. Princeton Paperbacks. Princeton University Press, Princeton, NJ, 1997.

[37] Sznitman, A-S: Topics in propagation of chaos. École d’Été de Probabilités de Saint-Flour XIX—1989, 165–251, Lecture Notes in Math., 1464, Springer, Berlin, 1991

[38] Villani, C: Topics in optimal transportation, Amer. Math. Soc., Providence, RI, 2003

[39] Villani, C: Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, Berlin, 2009

[40] Wang, X; Zhu, X: Kähler–Ricci solitons on toric manifolds with positive first Chern class, Advances in Mathematics 188 (2004), 87–103.

E-mail address: robertb@chalmers.se

Current address: Mathematical Sciences - Chalmers University of Technology and University of Gothenburg - SE-412 96 Gothenburg, Sweden