

Mean Field Approximations via Log-Concavity

Daniel Lacker¹, Sumit Mukherjee², and Lane Chun Yeung*

¹Department of Industrial Engineering & Operations Research, Columbia University, 500 W 120th St, NY 10027, USA

²Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA

*Corresponding author: Lane Chun Yeung. Email: l.yeung@columbia.edu

Communicated by Prof. Justin Solomon

We propose a new approach to deriving quantitative mean field approximations for any probability measure P on $\mathbb{R}^n$ with density proportional to $e^{f(x)}$, for f strongly concave. We bound the mean field approximation for the log partition function $\log \int e^{f(x)} dx$ in terms of $\sum_{i \neq j} \mathbb{E}_{Q^*} |\partial_{ij} f|^2$, for a semi-explicit probability measure Q^* characterized as the unique mean field optimizer, or equivalently as the minimizer of the relative entropy $H(\cdot | P)$ over product measures. This notably does not involve metric-entropy or gradient-complexity concepts which are common in prior work on nonlinear large deviations. Three implications are discussed, in the contexts of continuous Gibbs measures on large graphs, high-dimensional Bayesian linear regression, and the construction of decentralized near-optimizers in high-dimensional stochastic control problems. Our arguments are based primarily on functional inequalities and the notion of displacement convexity from optimal transport.

1 Introduction

At the center of the recent theory of *nonlinear large deviations* is the problem of justifying the *mean field approximation* for the partition function of a Gibbs measure. Given a (reference) probability measure ρ on $\mathbb{R}^n$, suppose a probability measure P on $\mathbb{R}^n$ takes the form

$$P(dx) = Z^{-1} e^{f(x)} \rho(dx),$$

for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and normalizing constant Z . A recurring problem in diverse applications is the approximation of the often intractable partition function Z . It obeys the well-known Gibbs variational principle (see Theorem 3.3 below)

$$\log Z = \log \int_{\mathbb{R}^n} e^f d\mu^{\otimes n} = \sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q | \rho) \right), \quad (1.1)$$

where $\mathcal{P}(\mathbb{R}^n)$ is the set of probability measures on $\mathbb{R}^n$, and H denotes the relative entropy

$$H(Q | Q') := \int_{\mathbb{R}^n} \frac{dQ}{dQ'} \log \frac{dQ}{dQ'} dQ' \text{ if } Q \ll Q', \quad H(Q | Q') := \infty \text{ if } Q \not\ll Q'.$$

Note that $Q = P$ is the unique optimizer in (1.1). Letting $\mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ denote the set of product measures $Q = Q_1 \times \cdots \times Q_n$ in $\mathcal{P}(\mathbb{R}^n)$, the mean field approximation is

$$\log \int_{\mathbb{R}^n} e^f d\mu^{\otimes n} \approx \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q | \rho) \right). \quad (1.2)$$

In the cases studied in this paper (see Section 2.1 for details) the left-hand side is expected to be of order n ; a precise formulation of (1.2) is then to find conditions under which the difference is $o(n)$, so that the mean field approximation becomes asymptotically correct at the leading order. Note that the right-hand side of (1.2) is trivially a lower bound for the left, because of (1.1), and it is only the upper bound that incurs an error, which must be estimated.

The groundbreaking work of [24], motivated by applications to subgraph counts in sparse random graphs, showed how to justify the mean field approximation in the case that ρ is the uniform measure on the hypercube $\{-1, 1\}^n$. Their key assumption is that the gradient of f has low complexity, as measured by the metric entropy of the range $\nabla f(\{-1, 1\}^n)$. A number of subsequent papers have since refined this approach and results on subgraph counts [25, 26, 49], in addition to other noteworthy applications such as Ising models [3, 4, 8, 28, 35, 44]. Most applications thus far involve discrete μ , but the theory has been extended to compactly supported measures [3, 5, 59]. Alternative and often more convenient estimates have appeared, still based on “gradient complexity” but quantifying it in a different way, eschewing covering number estimates in favor of the simpler and weaker Gaussian-width [35–37] or Rademacher-width [4].

In this paper, we propose an alternative approach to the mean field approximation, designed most notably for the case where f is concave and the reference measure μ is strongly log-concave (see Theorem 1.1 and Corollary 1.4). In particular, we deal with continuous μ of unbounded support, which covers a rather different host of applications, discussed in Section 2, compared to the somewhat more discrete-oriented prior literature. Our approach is based on a semi-explicit representation for the mean field optimizer Q^* in (1.2), which we show to be unique as soon as P is strictly log-concave, and which is in fact also the unique minimizer of $H(\cdot | P)$ over product measures. We control the error in the approximation (1.2) by a constant times $\mathbb{E}_{Q^*} \sum_{i \neq j} |\partial_{ij} f|^2$, which is typically much simpler to work with compared to the aforementioned notions of gradient complexity. Eldan [35, 36] and Austin [5] also analyze the mean field approximation by approximating P by product measures in entropy, but our methods and bounds are very different from theirs; notably, they approximate P not by a single product measure but by a mixture, which is natural when the mean field optimizer is not unique, as is explained well in [35]. The uniqueness of the mean field optimizer in our setting means that we expect P to concentrate around a single *pure state*, rather than a mixture of states.

In the rest of this section, we describe our general results on mean field approximations for log-concave measures, along with some related ideas and generalizations, with proofs deferred to Section 3. Section 2 develops three applications: Gibbs measures with heterogeneous interactions, high-dimensional Bayesian linear regression, and high-dimensional stochastic control problems.

1.1 Main results

Recall for $\kappa > 0$ that a function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ is said to be κ -concave if $x \mapsto f(x) + \frac{\kappa}{2}|x|^2$ is concave. If f is finite-valued and C^2 , that is, twice continuously differentiable, then f is κ -concave if and only if $\nabla^2 f(x) \leq -\kappa I$ in semidefinite order, for each $x \in \mathbb{R}^n$. We say that a probability measure P on $\mathbb{R}^n$ is κ -log-concave if it takes the form $P(dx) = e^{f(x)} dx$ for some κ -concave function f . We will work with the (negative of the) differential entropy

$$H(Q) := \int_{\mathbb{R}^n} Q(x) \log Q(x) dx,$$

for an absolutely continuous probability measure $Q(dx) = Q(x)dx$ on a Euclidean space, well-defined in $(-\infty, \infty]$ whenever the negative part of $Q \log Q$ is integrable; we adopt the convention that $H(Q) = \infty$ if Q is not absolutely continuous, or if $(Q \log Q)^-$ is not integrable. Let $\mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ denote the set of product measures on $\mathbb{R}^n$. Let $X = (X_1, \dots, X_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denote the identity map, so that we may write $\mathbb{E}_Q[g(X)] = \int_{\mathbb{R}^n} g dQ$ for the expectation under Q .

Theorem 1.1. Consider a C^2 and κ -log-concave probability measure $P(dx) = Z^{-1}e^{f(x)}dx$, for some $\kappa > 0$. Assume there exist $c_1 \geq 0$ and $0 \leq c_2 < \kappa/2$ such that $|f(x)| \leq c_1 e^{c_2|x|^2}$ for all $x \in \mathbb{R}^n$. Then the following conclusions hold:

- (1) There exists a unique product measure $Q^* = Q_1^* \times \cdots \times Q_n^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ with strictly positive density a.e. satisfying $f \in L^1(Q^*)$ and the fixed point equation

$$Q_i^*(dx_i) = Z_i^{-1} \exp(\mathbb{E}_{Q^*}[f(X) | X_i = x_i]) dx_i, \quad Z_i > 0, i = 1, \dots, n. \quad (1.3)$$

- (2) Q^* is κ -log-concave.
 (3) Q^* is the unique optimizer in

$$\sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right). \quad (1.4)$$

- (4) If we define

$$R_f := \log \int_{\mathbb{R}^n} e^{f(x)} dx - \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right),$$

then

$$0 \leq R_f \leq \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*}(\partial_i f(X) | X_i) \leq \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q^*}[|\partial_{ij} f(X)|^2]. \quad (1.5)$$

The supremum in (1.4) is finite, as we will see in Lemma 3.4. Also, as will be seen in the proof of Proposition 3.9, our assumptions ensure that $\partial f(x_i, \cdot) \in L^1(\prod_{j \neq i} Q_j^*)$, so the conditional variance in (1.5) is well-defined in $[0, \infty]$. The final quantity in (1.5) controlling our mean field approximation error involves only the cross-derivatives $i \neq j$, which are insensitive to additively separable perturbations $f(x) \rightarrow f(x) + \sum_{i=1}^n \tilde{f}_i(x_i)$. On the other hand, the measure Q^* is sensitive to these perturbations, but in the tractable sense that $Q_i^*(dx_i)$ must be multiplied by $\exp \tilde{f}_i(x_i)$ (and a new normalizing constant). In particular, both upper bounds in (1.5) vanish if f is already additively separable, that is, if P is a product measure.

In Theorem 1.1, the measure Q^* is defined implicitly, which can make bounding R_f difficult. In the simplest case where $\nabla^2 f$ is bounded, we need no knowledge of Q^* to obtain

$$R_f \leq \frac{1}{\kappa^2} \sup_{x \in \mathbb{R}^n} \sum_{1 \leq i < j \leq n} |\partial_{ij} f(x)|^2,$$

which is sharp enough for many applications. But even when $\nabla^2 f$ is unbounded, we can take advantage of the fact that Q^* is κ -log-concave by Theorem 1.1(2), which implies in particular that it has finite moments of all orders controlled in terms of κ .

A guiding example is the class of Gibbs measures with pairwise interactions of the form

$$f(x) = \sum_{i=1}^n V(x_i) + \sum_{1 \leq i < j \leq n} J_{ij} K(x_i - x_j), \quad (1.6)$$

where V is κ -concave, K is even and concave, and J is a symmetric matrix with nonnegative entries. Then $\partial_{ij} f(x) = -J_{ij} K''(x_i - x_j)$ for $i \neq j$, and for K'' bounded we immediately deduce $R_f \leq \text{Tr}(J^2) \|K''\|_\infty^2 / 2\kappa^2$ from Theorem 1.1. Corollary 2.3 below proves a similar $O(\text{Tr}(J^2))$ bound merely assuming that K'' has at most exponential growth, plus a symmetry assumption. Since $\log \int_{\mathbb{R}^n} e^f dx$ is order n in this case, we obtain a successful mean field approximation whenever J satisfies $\text{Tr}(J^2) = o(n)$, which is, in a sense, optimal. For instance, in the noteworthy case that J is $1/d$ times the adjacency matrix of a d -regular graph, we have $\text{Tr}(J^2) = o(n)$ precisely when $d \rightarrow \infty$, and the mean field approximation fails in general in the sparsest case where $d = O(1)$. See Section 2.1 for further discussion.

As a first corollary of Theorem 1.1, we deduce the following non-asymptotic law of large numbers for the empirical measure. We stress that this result is made possible by the uniqueness of Q^* in Theorem 1.1, and without uniqueness it can happen that the empirical measure remains random in the limit, as in the famous Curie-Weiss model; see [38] for the continuous-spin version, which violates our concavity assumption, and violates uniqueness at low temperature.

Corollary 1.2. Under the assumptions of Theorem 1.1, for any 1-Lipschitz function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, we have

$$\mathbb{E}_P \left[\left(\frac{1}{n} \sum_{i=1}^n \varphi(X_i) - \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{Q^*} [\varphi(X_i)] \right)^2 \right] \leq \frac{(1 + \sqrt{2R_f})^2}{\kappa n}. \quad (1.7)$$

Remark 1.3. Corollary 1.2 can be interpreted as a form of concentration of the empirical measure $\frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ around the measure $\frac{1}{n} \sum_{i=1}^n Q_i^*$. Alternatively, the Poincaré inequality for P implies $\text{Var}_P(\frac{1}{n} \sum_{i=1}^n \varphi(X_i)) \leq 1/\kappa n$ for 1-Lipschitz φ , which in turn implies a form of concentration of $\frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ around its mean $\frac{1}{n} \sum_{i=1}^n P_i$, where P_i is the i^{th} marginal of P . However, the latter is normally not as useful, because the marginals of P are typically not as tractable as the various characterizations of Q^* provided by Theorem 1.1. A concrete example is given by (1.6) with a doubly stochastic interaction matrix J , developed in detail in Theorem 2.5; the measures Q_i^* turn out to be *universal*, in the sense that they do not depend on i , n , or the choice of doubly stochastic matrix J , whereas the marginals of P may depend on all of these quantities.

It is often convenient to work with a probability measure as a reference measure, in place of Lebesgue measure, as is common in the literature on mean field approximations (see, e.g., [3, 5, 24, 35, 59]). Theorem 1.1 implies a similar result in terms of reference probability measures.

Corollary 1.4. Let $V_i : \mathbb{R} \rightarrow \mathbb{R}$ be C^2 and κ -concave for some $\kappa > 0$, such that $\rho_i(dx) = e^{V_i(x)} dx$ is a probability measure, for $i = 1, \dots, n$. Let $\rho = \rho_1 \times \dots \times \rho_n$. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be C^2 and concave. Assume there exist $c_1 \geq 0$ and $0 \leq c_2 < \kappa/2$ such that $|g(x)| \leq c_1 e^{c_2 |x|^2}$ for all $x \in \mathbb{R}^n$. Then the following conclusions hold:

- (1) There exists a unique product measure $Q^* = Q_1^* \times \dots \times Q_n^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ with strictly positive density a.e. satisfying $g \in L^1(Q^*)$ and

$$Q_i^*(dx_i) = Z_i^{-1} \exp(\mathbb{E}_{Q^*}[g(X) | X_i = x_i]) \rho_i(dx_i), \quad Z_i > 0, \quad i = 1, \dots, n. \quad (1.8)$$

- (2) Q^* is κ -log-concave.
 (3) Q^* is the unique optimizer in

$$\sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - H(Q | \rho) \right). \quad (1.9)$$

- (4) If we define

$$R_g^\rho := \log \int_{\mathbb{R}^n} e^g d\rho - \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - H(Q | \rho) \right),$$

then

$$0 \leq R_g^\rho \leq \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*}(\partial_i g(X) | X_i) \leq \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q^*} [|\partial_{ij} g(X)|^2]. \quad (1.10)$$

For certain symmetric choices of g , the bound (1.10) is related to the theorems of Cramér and Sanov on large deviations, which are settings in which the Gibbs variational principle is well known to be nearly

saturated by product measures. For instance, if $g(x) = nG(\frac{1}{n} \sum_{k=1}^n x_k)$ for some continuous concave G , we obtain $R_g^\rho \leq \|G''\|_\infty^2 / 2\kappa^2$, which is certainly $o(n)$ when G'' is bounded.

1.2 Overview and proof ideas

We explain here some key ideas behind Theorem 1.1 and its corollaries. The simple identity

$$\log \int_{\mathbb{R}^n} e^{f(x)} dx - \int_{\mathbb{R}^n} f dQ + H(Q) = H(Q|P) \quad (1.11)$$

is valid for probability measures Q with finite entropy and implies (see Lemma 3.4 for details)

$$\log \int_{\mathbb{R}^n} e^{f(x)} dx - \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right) = \inf_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} H(Q|P), \quad (1.12)$$

and also that optimizing (1.4) is equivalent to optimizing

$$\inf_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} H(Q|P). \quad (1.13)$$

That is, Q^* from Theorem 1.1 is the optimizer in (1.13). This can be seen as an *entropic projection*, in the sense of Csiszar [27], onto the set of product measures. A minimizer in (1.13) always exists, because the set of product measures is weakly closed and $H(\cdot|P)$ has weakly compact sub-level sets. But uniqueness is not obvious and in fact fails in general, because the set of product measures is not convex. We establish the uniqueness of the optimizer in Lemma 3.6 in the case where P is strictly log-concave, by exploiting the notion of displacement convexity from the theory of optimal transport, with similarities to the work of McCann [50].

Once we know that the optimizer Q^* for (1.4) takes the form (1.3), the proof of the mean field approximation (1.5) is fairly quick, if we ignore certain technical points: The right-hand side of the identity (1.12) is precisely $H(Q^*|P)$. We first use the log-Sobolev inequality for P , which is ensured by κ -log-concavity and the famous result of Bakry-Émery [6], to get

$$H(Q^*|P) \leq \frac{1}{2\kappa} \int_{\mathbb{R}^n} \left| \nabla \log \frac{dQ^*}{dP} \right|^2 dQ^*.$$

Since $Q^* = Q_1^* \times \cdots \times Q_n^*$ is a product measure, the formula (1.3) implies

$$\partial_i \log Q^*(x) = \partial_i \log Q_i^*(x_i) = \partial_i \mathbb{E}_{Q^*} [f(X) | X_i = x_i] = \mathbb{E}_{Q^*} [\partial_i f(X) | X_i = x_i]. \quad (1.14)$$

Thus,

$$H(Q^*|P) \leq \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \left(\mathbb{E}_{Q^*} [\partial_i f(X) | X_i] - \partial_i f(X) \right)^2 = \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*} (\partial_i f(X) | X_i).$$

Differentiating (1.14) again shows easily that Q^* is κ -log-concave since f is concave. Hence, Q^* and its marginals obey a Poincaré inequality, and we deduce

$$\text{Var}_{Q^*} (\partial_i f(X) | X_i) \leq \frac{1}{\kappa} \sum_{j \neq i} \mathbb{E}_{Q^*} [|\partial_{ij} f(X)|^2 | X_i].$$

Combining the last two inequalities yields (1.5). See Section 1.3.3 below for a discussion of a generalization of this argument beyond the strongly log-concave case.

The proof of Corollary 1.2 begins with the observation that the κ -log-concavity of P in Theorem 1.1 implies the quadratic transport inequality [53, Theorems 1 and 2]

$$\mathcal{W}_2^2(Q^*, P) \leq \frac{2}{\kappa} H(Q^*|P), \quad (1.15)$$

where $\mathcal{W}_2$ denotes the quadratic Wasserstein distance defined by

$$\mathcal{W}_2^2(Q^*, P) = \inf_{\pi} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \pi(dx, dy),$$

where the infimum is over $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ with marginals Q^* and P . Combining (1.15) with the inequality $H(Q^* | P) = R_f$ discussed above, we arrive at $\mathcal{W}_2^2(Q^*, P) \leq 2R_f/\kappa$. The quadratic Wasserstein distance enjoys a useful and fairly well known *subadditivity* inequality, which we prove in Section 3.4 for the sake of completeness: If P_S denotes the marginal law of $(X_i)_{i \in S}$ under P for a set $S \subset [n] := \{1, \dots, n\}$, and similarly for Q_S^* , then we have

$$\binom{n}{k}^{-1} \sum_{S \subset [n], |S|=k} \mathcal{W}_2^2(Q_S^*, P_S) \leq \frac{1}{[n/k]} \mathcal{W}_2^2(Q^*, P) \leq \frac{2}{\kappa [n/k]} R_f \leq \frac{4k}{n\kappa} R_f \quad (1.16)$$

for any $1 \leq k \leq n$. With (1.16) in hand, the proof of Corollary 1.2 is straightforward by Kantorovich duality and Poincaré inequality (see Section 3.4 for a detailed proof). Moreover, in our cases of interest where $R_f = o(n)$, the bound (1.16) quantifies a form of *approximate independence*: Most k -particle marginals of P are $\mathcal{W}_2$ -close to product measures, if $k = o(n/R_f)$.

Remark 1.5. We work throughout the paper with state space $\mathbb{R}$, for simplicity. That is, we study approximations of measures on $\mathbb{R}^n$ by n -fold products of measures on $\mathbb{R}$, as opposed to, say, approximations of measures on $(\mathbb{R}^d)^n$ by n -fold products of measures on $\mathbb{R}^d$. Most of our arguments, based primarily on convexity and functional inequalities, extend to the case of $\mathbb{R}^d$ or even Riemannian manifolds with lower curvature bounds in the spirit of Bakry-Émery [6, 7]. The only difficulty is in the uniqueness claimed in Theorem 1.1 (proven in Proposition 3.9), which would require a finer analysis involving regularity of certain optimal transport maps.

1.3 Additional discussion and results

The remaining results presented in this section will not be used in the rest of the paper but serve to elaborate on the structure of the main theorem. The reader mainly interested in applications or proofs of the above results may skip to Sections 2 or 3, respectively, with no loss of continuity.

1.3.1 More on entropic projections

Reversing the order of arguments in the relative entropy in (1.13) leads to a very different optimization problem, but it is instructive to compare the two. The infimum

$$\inf_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} H(P | Q) \quad (1.17)$$

is uniquely attained by taking $Q = P^* := P_1 \times \dots \times P_n$ to be the product of the marginals of P . Indeed, from the simple identity $H(P | Q) = H(P | P^*) + H(P^* | Q)$, it follows that $H(P | Q) \geq H(P | P^*)$ for all Q , with equality if and only if $Q = P^*$.

The Gaussian case highlights the difference between (1.17) and (1.13). Suppose P is a centered Gaussian with nonsingular covariance matrix Σ . In this case it is easy to see that the (unique) minimizer of $H(Q | P)$ among product measures Q is the centered Gaussian with covariance matrix $\hat{\Sigma}$, where $\hat{\Sigma}^{-1}$ is the diagonal matrix obtained by deleting the off-diagonal entries of Σ^{-1} . On the other hand, the unique minimizer of $H(P | Q)$ among product measures Q is the centered Gaussian with covariance matrix $\hat{\Sigma}$ obtained by deleting the off-diagonal entries of Σ .

1.3.2 Tilts

A similar bound to Corollary 1.4 is available if one seeks a stronger mean field approximation, in which $\mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ is replaced by the sub-class of product measures given by tilts of a given reference measure. We focus on the case of Gaussian reference measure, as it is not obvious how to extend the argument to a general reference measure. For $y \in \mathbb{R}^n$, let $\gamma_{y,t}$ denote the Gaussian with mean y and covariance matrix tI , with $\gamma_t := \gamma_{0,t}$, noting that $\gamma_{y,t} \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$.

Proposition 1.6. Let $t > 0$, and let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be C^2 and concave. Assume there exist $c_1 \geq 0$ and $0 \leq c_2 < 1/2t$ such that $|f(x)| \leq c_1 e^{c_2|x|^2}$. Then there is a unique $y^* \in \mathbb{R}^n$ satisfying

$$y^* = t \int_{\mathbb{R}^n} \nabla f d\gamma_{y^*,t}, \quad (1.18)$$

and it holds that

$$\log \int_{\mathbb{R}^n} e^f d\gamma_t \leq \sup_{y \in \mathbb{R}^n} \left(\int_{\mathbb{R}^n} f d\gamma_{y,t} - H(\gamma_{y,t} | \gamma_t) \right) + \frac{t^2}{2} \sum_{i,j=1}^n \int_{\mathbb{R}^n} |\partial_{ij} f|^2 d\gamma_{y^*,t}. \quad (1.19)$$

Noting that $H(\gamma_{y,t} | \gamma_t) = |y|^2/2t$, a simple calculation shows that y^* uniquely attains the supremum in (1.19). The difference between (1.19) and (1.10) is that the former includes the diagonal terms $i = j$ in the sum. This is natural; an additively separable function $f(x) = \sum_{i=1}^n f_i(x_i)$ yields a product measure $P(dx) = Z^{-1} e^{f(x)} \gamma_t(dx)$, but it takes an *affine* function f for P to be a Gaussian. Small off-diagonal derivatives $\partial_{ij} f$ can be naturally interpreted as meaning f is close to being additively separable, but the full Hessian matrix $\nabla^2 f$ must to be small in order for f to be close to affine.

The above proposition is worth comparing with prior results based on gradient complexity. It was shown in [4, Proposition 3.4, arXiv version] that if $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is C^1 then

$$\log \int_{\mathbb{R}^n} e^f d\gamma_t \leq \sup_{y \in \mathbb{R}^n} \left(\int_{\mathbb{R}^n} f d\gamma_{y,t} - H(\gamma_{y,t} | \gamma_t) \right) + \sqrt{2} \int_{\mathbb{R}^n} \sup_{y \in \mathbb{R}^n} (x \cdot \nabla f(y)) \gamma_t(dx). \quad (1.20)$$

The last integral is $(\sqrt{t})$ times the *Gaussian mean-width* of the set $\nabla f(\mathbb{R}^n)$. This estimate (1.20) has the advantage of applying to non-concave functions f , but it is only meaningful if ∇f is bounded. Proposition 1.6, on the other hand, can accommodate non-Lipschitz but concave functions f .

1.3.3 Generalization of the main theorem

We briefly discuss how Theorem 1.1 can generalize beyond the strongly log-concave setting. Essentially, strong log-concavity is needed only for the uniqueness claims and to justify the log-Sobolev and Poincaré inequalities as explained in Section 1.2. *Uniqueness* of Q^* is actually not essential if one is interested only in a bound like (1.5). The *existence* of an optimizer Q^* is automatic, and it is not hard to show that it must satisfy the fixed point equation (1.3), modulo technical conditions. If it can be shown that Q^* admits a strictly positive C^2 density, and that P and Q^* obey a log-Sobolev and Poincaré inequality, respectively, with constants C_1 and C_2 , then the following bound can be proven as in Section 1.2:

$$0 \leq R_f \leq C_1 \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*}(\partial_i f(X) | X_i) \leq 2C_1 C_2 \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q^*} [|\partial_{ij} f(X)|^2].$$

It is unclear if our assumed bound on $|f(x)|$ is needed or merely an artifact of our proof technique. We use the assumed bound on $|f(x)|$ in the proof of Theorem 1.1 only to show that Q^* is strictly positive a.e., but this can be shown directly in many particular cases, such as when f is symmetric.

1.4 Outline of the paper

In Section 2, we will present in detail the three main applications of Theorem 1.1, which pertain to Gibbs measures, high-dimensional Bayesian linear regression, and high-dimensional stochastic optimal control. The proof of Theorem 1.1 is given in Section 3.1, followed by the proof of Corollary 1.4 in Section 3.2. Section 3.3 contains the proof of Proposition 1.6, while Section 3.4 contains the proofs of the subadditivity inequality (1.16) and Corollary 1.2. Finally, the proofs of the applications are given in Sections 4 and 5.

2 Applications

2.1 Gibbs measures with pairwise interactions

First, we study Gibbs measures with pairwise interaction potentials of the form (1.6), where the following assumption holds:

Assumption 2.1. $V : \mathbb{R} \rightarrow \mathbb{R}$ is C^2 and κ -concave for some $\kappa > 0$, $K : \mathbb{R} \rightarrow \mathbb{R}$ is even, C^2 , and concave, and J is a symmetric matrix with nonnegative entries and $J_{ii} = 0$ for all $i = 1, \dots, n$. Assume there exists $a, b, c \geq 0$ and $0 \leq d < \kappa/2$ such that $|V(x)| \leq ce^{dx^2}$ and $|K''(x)|^2 \leq ae^{b|x|}$ hold for all $x \in \mathbb{R}$.

Note since K is even that there is no loss of generality in assuming that J is zero on the diagonal. The most traditional *mean field* setting is when $J_{ij} = 1/n$ for all (i, j) , so that all particles interact equally, and there is a vast literature on the large- n behavior; see [23, 34] for some recent results and references. In general, the matrix J represents *disorder* or *heterogeneous interactions*, and a common situation is when J is the rescaled adjacency matrix of a graph. A notable strength of the non-asymptotic perspective of our work, and the theory of nonlinear large deviations more broadly, is that it can seamlessly handle this kind of heterogeneity. Gibbs measures with pairwise interactions on large graphs have been studied in many contexts, primarily on finite state space (see [8, 10, 31, 32] and references therein). In the continuous context we study here, these Gibbs measures appear as invariant measures of locally interacting diffusion processes whose large-scale behavior has recently been the subject of active research [29, 52]. We note that concavity is not an uncommon assumption in this literature, and covers for instance models of granular media [22].

To work toward applying Theorem 1.1 with f as in (1.6), we first record the simple observation that f is strongly concave under Assumption 2.1. The proof of this and other results in Section 2.1 are given in Section 4.

Lemma 2.2. Define f by (1.6), and suppose Assumption 2.1 holds. Then f is κ -concave.

The following corollary will allow us to cover the case of unbounded K'' , but only if we can control the barycenter of Q^* in the sense that $\mathbb{E}_{Q^*}[X_i - X_j] = 0$. This symmetry condition is justified in different ways in the following applications and is explained further in Section 2.1.3.

Corollary 2.3. Define f by (1.6), and suppose Assumption 2.1 holds. With Q^* denoting the unique optimizer of (1.4), assume further that $\mathbb{E}_{Q^*}[X_i - X_j] = 0$. Then

$$R_f \leq \text{Tr}(J^2) a \kappa^{-2} e^{b^2/\kappa}.$$

Remark 2.4. Corollary 2.3 shows that $R_f = o(n)$ as long as $\text{Tr}(J^2) = o(n)$. The assumption $\text{Tr}(J^2) = o(n)$ has been used in the literature as a *mean field condition* for quadratic interaction models, first in [8, Theorem 1.1] and then in [59, Theorem 4]. Both cases are limited to measures with compact support. Moreover, in their setting, neither uniqueness of the optimizer nor convergence of the empirical measure hold in general. In contrast, the powerful assumption of concavity in our setting allows us to handle unbounded support and to show both uniqueness of the optimizer and convergence of the empirical measure Theorems 2.5 and 2.8 below.

Using Corollary 2.3, one can study the weak law of large numbers of the empirical measure under P , by studying the corresponding weak law under the product measure Q^* . Under additional assumptions on the matrix J , the mean field optimization problem can be shown to converge as $n \rightarrow \infty$, allowing us to characterize the weak law under P in terms of the limiting optimization problem. Below we illustrate this in two special cases.

2.1.1 Doubly stochastic matrices

In the following $n \rightarrow \infty$ results, note that the dependence of f , $P(dx) = Z^{-1}e^{f(x)}dx$ and J on n is suppressed.

Theorem 2.5. Define f by (1.6), and suppose Assumption 2.1 holds. Assume there exist $a, b \geq 0$ such that $|K''(x)|^2 \leq ae^{b|x|}$ for all x . Assume further that the symmetric matrix J is doubly stochastic (i.e., $\sum_{j=1}^n J_{ij} = 1$ for all i), and obeys the mean field condition $\text{Tr}(J^2) = o(n)$. Then we have the following conclusions:

(1)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{R}^n} e^{f(x)} dx = \sup_{Q \in \mathcal{P}(\mathbb{R})} \left(\int_{\mathbb{R}} V dQ + \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q(dx) Q(dy) - H(Q) \right). \quad (2.1)$$

(2) The supremum in (2.1) is attained by a unique $Q \in \mathcal{P}(\mathbb{R})$, and if $(X_1, \dots, X_n) \sim P$ then

$$\frac{1}{n} \sum_{i=1}^n \delta_{X_i} \rightarrow Q, \text{ weakly in law.} \quad (2.2)$$

(3) The measure Q^* of Theorem 1.1 is precisely $Q^* = Q^{\otimes n}$, and in particular it does not depend on n or the choice of doubly stochastic interaction matrix J .

The above theorem applies when $J = A/d$ and A is the adjacency matrix of a d -regular graph. In this case we get $\text{Tr}(J^2) = n/d$, which is $o(n)$ as long as $d \rightarrow \infty$. The above theorem is similar in spirit to [8, Theorem 2.1], which dealt with Ising and Potts models, and a comment similar to Remark 2.4 applies. Note that one cannot expect a mean field approximation to be valid in the sparsest (diluted) case, where d stays bounded as $n \rightarrow \infty$. The framework of local weak convergence has proven to be successful in this context [30], and we refer also to [45, Sections 2 and B] for continuous models encompassing the form studied here, and for a detailed derivation of the (folklore) limit of the empirical measure for locally convergent graph sequences, which requires uniqueness of the infinite-volume Gibbs measure on the limiting graph.

2.1.2 Graphons

Another case in which we can derive asymptotics of the log partition function is when the matrix J converges to a graphon W in cut metric. Below we introduce the relevant notions, deferring to [15–17, 48] for additional background:

Definition 2.6. Let $\mathcal{W}$ denote the space of all symmetric measurable functions from $[0, 1]^2$ to $[0, \infty)$ which are integrable. For $W_1, W_2 \in \mathcal{W}$, define the strong cut (pseudo-)metric by

$$d_{\square}(W_1, W_2) := \sup_{S, T \subset [0, 1]} \left| \int_{S \times T} (W_1(u, v) - W_2(u, v)) du dv \right|,$$

and their weak cut (pseudo-)metric by

$$\delta_{\square}(W_1, W_2) := \inf_{\varphi} d_{\square}(W_1, W_2^{\varphi}),$$

where the infimum is over all invertible measure-preserving maps $\varphi : [0, 1] \rightarrow [0, 1]$, and $W_2^{\varphi}(u, v) := W_2(\varphi(u), \varphi(v))$. Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with nonnegative entries, we define a function $W_A \in \mathcal{W}$ by setting $W_A(u, v) := A_{\lceil nu \rceil, \lceil nv \rceil}$. We say that a sequence of symmetric matrices $\{A_n\}$ converges in weak cut metric to a function $W \in \mathcal{W}$ if $\delta_{\square}(W_{A_n}, W) \rightarrow 0$.

Remark 2.7. Suppose G_n is the adjacency matrix of an Erdős-Rényi random graph on n vertices with parameter p_n , such that $np_n \rightarrow \infty$. If $J_n = \frac{1}{np_n} G_n$, then nJ_n converges in strong cut metric to the constant function 1 (see [15, Example 3.3.1]). Similar convergences hold if G_n arises from a stochastic block model, where the edge probability matrix has a block structure, in which case the limiting W retains the same block structure. For more examples of convergent sequence of graphs in cut metric, we refer again to [15–17, 48] and references therein.

Let $\mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ denote the space of all probability measures on $[0, 1] \times \mathbb{R}$ with uniform first marginal. Note that any $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ admits the disintegration $\mu(du, dx) = du\mu_u(dx)$.

Theorem 2.8. Define f by (1.6), and suppose Assumption 2.1 holds. Assume there exist $a, b \geq 0$ such that $|K''(x)|^2 \leq ae^{b|x|}$ for all x . Assume further that V is even, K is nonpositive, $\int_{\mathbb{R}} e^{V(x)} dx = 1$, and $J = \{J_n\}$ is a sequence of matrices such that $\{J_n\}$ converges in weak cut metric to a function $W \in \mathcal{W}$. Assume also that $\text{Tr}(J_n^2) = o(n)$.

(1) Defining the probability measure $\rho(dx) = e^{V(x)} dx$, we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{R}^n} e^{f(x)} dx \\ &= \sup_{\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})} \left(\frac{1}{2} \int_{([0, 1] \times \mathbb{R})^2} W(u, v) K(x - y) \mu(du, dx) \mu(dv, dy) - \int_0^1 H(\mu_u | \rho) du \right). \end{aligned} \quad (2.3)$$

(2) The supremum in (2.3) is attained by a unique $\mu^* \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$, and if $(X_1, \dots, X_n) \sim P$, then

$$\frac{1}{n} \sum_{i=1}^n \delta_{X_i} \rightarrow \int_0^1 \mu_u^* du, \text{ weakly in law.} \quad (2.4)$$

Remark 2.9. It follows from [16, Propositions C.5 and C.15] that the condition $\text{Tr}(J_n^2) = o(n)$ holds automatically if J_n is the adjacency matrix of a simple graph $G_n = ([n], E_n)$ multiplied by $n/(2|E_n|)$, and nJ_n converges in cut metric. However, if J_n is a general matrix, we need the added assumption $\text{Tr}(J_n^2) = o(n)$ in Theorem 2.8.

2.1.3 On the symmetry of Q^*

This short section elaborates on conditions under which one can check that $\mathbb{E}_{Q^*}[X_i - X_j] = 0$, which was needed in Corollary 2.3. The main two conditions we found are evenness and a weak form of permutation invariance.

Definition 2.10. Let S be a set of permutations of $[n]$. We say that S is *transitive* if for every $i, j \in [n]$ there exists $\pi \in S$ such that $\pi(i) = j$. We say also that a function f on $\mathbb{R}^n$ is *invariant under S* if $f(X_1, \dots, X_n) = f(X_{\pi(1)}, \dots, X_{\pi(n)})$ for every $x \in \mathbb{R}^n$ and $\pi \in S$.

Lemma 2.11. In the setting of Theorem 1.1, the following implications hold:

- (1) If f is even, meaning $f(-x) = f(x)$ for all x , then Q_i^* is even for each $i = 1, \dots, n$.
- (2) Suppose f is invariant under a transitive set of permutations. Then $Q_1^* = Q_2^* = \dots = Q_n^*$.

In both cases, we have $\mathbb{E}_{Q^*}[X_i - X_j] = 0$ for all $i, j \in [n]$.

When f is of the form (1.6), it is clear that f is even if K and V are, and indeed V is assumed even in Theorem 2.5 to enable an application of Lemma 2.11(1). We will not apply Lemma 2.11(2), but we find it interesting in its own right. For instance, (2) holds if f is symmetric, that is, invariant under all permutations. Another natural case covered by (2) is where f is of the form (1.6) and J is a scalar multiple of the adjacency matrix of a *vertex transitive* graph.

2.2 High dimensional Bayesian linear regression

Our next application is concerned with high dimensional Bayesian linear regression. Suppose we observe a set of data $\{(y_i, X_i)\}_{i=1}^n$, where $y_i \in \mathbb{R}$ and $X_i \in \mathbb{R}^p$. Let $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$ and $\mathbf{X}^T = (X_1, \dots, X_n) \in \mathbb{R}^{p \times n}$. Consider the linear regression model

$$y = \mathbf{X}\beta + \varepsilon, \quad \varepsilon \sim \gamma_{\sigma^2},$$

where γ_{σ^2} denotes the Gaussian with mean 0 and covariance matrix $\sigma^2 I$. Here $\beta \in \mathbb{R}^p$ is the unknown parameter.

Following a Bayesian approach, assume that $\beta = (\beta_1, \dots, \beta_p)^\top \stackrel{\text{i.i.d.}}{\sim} \pi$, where π is a prior distribution on $\mathbb{R}$ with density proportional to $e^V \in L^1(\mathbb{R})$ for some $V : \mathbb{R} \rightarrow \mathbb{R}$. The posterior density $\pi_{y,\mathbf{x}}$ of β given y and $\mathbf{x}$ is then proportional to $e^{f_{y,\mathbf{x}}}$, where

$$f_{y,\mathbf{x}}(\beta) := \sum_{i=1}^p V(\beta_i) - \frac{1}{2\sigma^2} |y - \mathbf{x}\beta|^2.$$

The posterior distribution is the central object of inference in Bayesian statistics. Note that even though β has independent coordinates under the prior, the coordinates of β are no longer independent under the posterior. Frequently, mean-field techniques are used to approximate such complex posterior distributions, including and beyond the set up of Bayesian linear regression (see [2, 11, 54, 58, 60] and references therein). In particular, it is useful to understand what conditions guarantee the validity of a mean field approximation, showing that the posterior is close to a product measure. Using Theorem 1.1, the following corollary provides sufficient conditions under which the posterior is indeed mean-field. Leveraging this, it also derives a law of large numbers for the empirical measure under the true posterior distribution.

Corollary 2.12. Assume V is κ_1 -concave for some $\kappa_1 \in \mathbb{R}$, and that there exists $c_1 \geq 0$ and $0 \leq c_2 < \kappa/2$ such that $|V(x)| \leq c_1 e^{c_2 x^2}$ for all $x \in \mathbb{R}$. Set $J = \mathbf{X}^\top \mathbf{X} \in \mathbb{R}^{p \times p}$, and assume that $J \geq \kappa_2 I$ for some $\kappa_2 \in \mathbb{R}$ such that $\kappa_1 + \kappa_2 \sigma^{-2} > 0$. Then

$$\sup_{y \in \mathbb{R}^n} \left| \log \int_{\mathbb{R}^p} e^{f_{y,\mathbf{x}}(\beta)} d\beta - \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^p)} \left(\int_{\mathbb{R}^p} f_{y,\mathbf{x}} dQ - H(Q) \right) \right| \leq \frac{1}{(\kappa_1 \sigma^2 + \kappa_2)^2} \sum_{1 \leq i < j \leq p} J_{ij}^2. \quad (2.5)$$

Moreover, for every $y \in \mathbb{R}^n$, the inner supremum in (2.5) is attained by a unique $Q_y^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^p)$, and for any 1-Lipschitz function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, we have

$$\sup_{y \in \mathbb{R}^n} \mathbb{E}_{\pi_{y,\mathbf{x}}} \left[\left(\frac{1}{p} \sum_{i=1}^p \varphi(\beta_i) - \frac{1}{p} \sum_{i=1}^p \mathbb{E}_{Q_y^*}[\varphi(\beta_i)] \right)^2 \right] \leq \frac{\sigma^2 (\kappa_1 \sigma^2 + \kappa_2 + \sqrt{2 \sum_{1 \leq i < j \leq p} J_{ij}^2})^2}{p(\kappa_1 \sigma^2 + \kappa_2)^3}. \quad (2.6)$$

The proof of this corollary is by a direct application of Theorem 1.1 and Corollary 1.2, and is hence omitted. Indeed, the concavity assumption on V and the lower bound on J ensure that $\nabla^2 f_{y,\mathbf{x}}(\beta) \leq -(\kappa_1 + \kappa_2 \sigma^{-2})I$ for all β .

Remark 2.13. The uniformity in y in (2.5) implies that the mean field approximation continues to hold with high probability, under any distributional assumption on y . Note that when $n, p \rightarrow \infty$ in any arbitrary manner, the right-hand side of (2.5) and (2.6) are $o(p)$ and $o(1)$ respectively, as long as $\sum_{1 \leq i < j \leq p} J_{ij}^2 = o(p)$ when $n, p \rightarrow \infty$. We also point out that the same conclusion as in (2.5) above was derived in [51, Theorem 1] using very different techniques, under the assumption that the prior distribution π is compactly supported. In our setup, we allow the support to be non-compact and instead assume that the prior distribution is strongly log-concave. This powerful assumption brings many advantages, such as yielding the uniqueness of the optimizer and the law of large numbers under no extra assumptions.

2.3 Stochastic control

This section describes an application of Corollary 1.4 to a class of high-dimensional stochastic optimal control problems. Let $T > 0$, and let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be C^2 and concave. Consider the stochastic control problem

$$V_{\text{orig}} := \sup \mathbb{E} \left[g(X_T) - \frac{1}{2n} \sum_{i=1}^n \int_0^T |\alpha_i(t, X_t)|^2 dt \right], \quad (2.7)$$

where the supremum is over pairs (α, X) , where $\alpha = (\alpha_1, \dots, \alpha_n) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a measurable function and $X = (X^1, \dots, X^n)$ a weak solution of the stochastic differential equation (SDE)

$$dX_t^i = \alpha_i(t, X_t)dt + dB_t^i, \quad X_0^i = 0, \quad i = 1, \dots, n, \quad (2.8)$$

defined on an arbitrary filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, satisfying also $\int_0^T |\alpha(t, X_t)|^2 dt < \infty$ a.s. Here $B = (B^1, \dots, B^n)$ is an n -dimensional $\mathbb{F}$ -Brownian motion, and X is required to be $\mathbb{F}$ -adapted. We call such a pair (α, X) *admissible*. There is a well known semi-explicit solution to (2.7) which has come to be known as the *Föllmer drift*, which we will discuss in Remark 2.15 below. Let us stress that concavity of g is, again, both a powerful assumption and yet quite commonplace in both theory and applications of stochastic optimal control [39].

We interpret $i = 1, \dots, n$ as the indices of different “players,” each facing an independent source of randomness B^i , and each choosing a control α_i which can depend on the full information of all n players. Players “cooperate” in the sense that $(\alpha_1, \dots, \alpha_n)$ are chosen together to optimize (2.7). When g is of the form

$$g(x) = G\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_i}\right), \quad \text{for some } G : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}, \quad (2.9)$$

we recover a well-studied class of problems which goes under the name *mean field control* in the cooperative setting [20], or *mean field games* in the competitive (Nash equilibrium) setting [42, 46]; see [21] for an overview. In this setting, it is typically argued that V_{orig} converges to the value of a limiting “mean field” control problem, and the optimal control $\hat{\alpha}$ from this limiting problem can be used to construct *distributed* controls $\alpha_i(t, x_1, \dots, x_n) = \hat{\alpha}(t, x_i)$ which are provably approximately optimal for the n -player problem for n large. This is a very desirable outcome, because distributed controls are much simpler (lower-dimensional).

Our results give a new non-asymptotic perspective on control problems of this form, by showing how to construct approximately optimal distributed controls for much more general g than in (2.9). The link between (2.7) and the setting of Section 1 is the formula

$$V_{\text{orig}} = \sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - \frac{1}{n} H(Q | \gamma_T) \right) = \frac{1}{n} \log \int_{\mathbb{R}^n} e^{ng} d\gamma_T, \quad (2.10)$$

where we recall that γ_T denotes the centered Gaussian with covariance matrix T . This formula is essentially a well known consequence of Girsanov’s theorem. (Experts might recognize a similarity with a famous formula often named after Boué-Dupuis [18] or Borell [14], though the form we present here is simpler because of our restriction to Markovian controls, whereas [14, 18] work with *open-loop* controls, that is, controls specified as arbitrary progressively measurable processes.) The mean field approximation also admits a natural control-theoretic interpretation. Define

$$V_{\text{dstr}} := \sup \mathbb{E} \left[g(X_T) - \frac{1}{2n} \sum_{i=1}^n \int_0^T |\alpha_i(t, X_t)|^2 dt \right], \quad (2.11)$$

where the supremum is now over admissible pairs (α, X) for which $\alpha = (\alpha_1, \dots, \alpha_n)$ is of the form

$$\alpha_i(t, x_1, \dots, x_n) = \hat{\alpha}_i(t, x_i),$$

for some measurable $\hat{\alpha}_i : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, and also for which $X_t^1, \dots, X_t^n$ are independent for each $t \in [0, T]$ (this second statement being redundant if the SDE (2.8) driven by this α is known to be unique in law). Let us call any such pair (α, X) a *distributed admissible pair*. We will derive the following result from Corollary 1.4, after first showing that V_{dstr} is nothing but the mean field approximation of (2.10), in the sense that

$$V_{\text{dstr}} = \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - \frac{1}{n} H(Q | \gamma_T) \right). \quad (2.12)$$

Corollary 2.14. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be C^2 and concave, and let $T > 0$. Assume there exists $c_1 \geq 0$ and $0 \leq c_2 < 1/2T$ such that $|g(x)| \leq c_1 e^{c_2|x|^2}$ for all $x \in \mathbb{R}^n$. Define V_{orig} and V_{dstr} by (2.7) and (2.11), respectively. Then the formulas (2.10) and (2.12) hold, and

$$0 \leq V_{\text{orig}} - V_{\text{dstr}} \leq nT^2 \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q^*} [|\partial_{ij} g(X)|^2], \quad (2.13)$$

where $Q^* = Q_1^* \times \cdots \times Q_n^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ is the unique product measure with strictly positive density a.e. satisfying $g \in L^1(Q^*)$ and the fixed point equation

$$Q_i^*(dx_i) = Z_i^{-1} \exp(n \mathbb{E}_{Q^*}[g(X) | X_i = x_i]) \gamma_T(dx_i), \quad Z_i > 0, \quad i = 1, \dots, n.$$

The proof is given in Section 5. Corollary 2.14 shows that distributed controls are approximately optimal for large n if $n \|\sum_{i \neq j} \partial_{ij} g\|_\infty^2 = o(1)$. As an example, if g is of the form (2.9) and G is twice continuously Wasserstein- or L -differentiable in the sense of [21, Chapter 5.2], then

$$\partial_i g(x) = \frac{1}{n} D_m G \left(\frac{1}{n} \sum_{k=1}^n \delta_{x_k}, x_i \right), \quad \partial_{ij} g(x) = \frac{1}{n^2} D_m^2 G \left(\frac{1}{n} \sum_{k=1}^n \delta_{x_k}, x_i, x_j \right), \quad i \neq j.$$

Hence, if $D_m^2 G$ is bounded, then the right-hand side of (2.13) is bounded by $T^2 \|D_m^2 G\|_\infty^2 / 2n$.

Remark 2.15. In fact, the proof of Corollary 2.14 also yields an explicit characterization of the optimal distributed control in (2.11), which we summarize as follows. For a measure $Q \ll \gamma_T$, consider a process $X = (X_t)_{t \in [0, T]}$ such that $X_T \sim Q$ and the conditional law of the trajectory $(X_t)_{t \in [0, T]}$ given $X_T = x$ coincides with the law of the Brownian bridge from 0 to x on the time interval $[0, T]$. This process might be called the *Brownian (or Schrödinger) bridge with terminal law Q* . The associated control α is given by $\alpha(t, x) = \nabla_x \log \mathbb{E}[\frac{dQ}{d\gamma_T}(x + B_T - B_t)]$, as shown in full generality by Föllmer [40, 41]. Note that the associated SDE (2.8) may not be pathwise unique in general, but it always admits a weak solution X with the law just described. The optimizer for the original control problem (2.7) is nothing but the Brownian bridge with terminal law $P(dx) = Z^{-1} e^{ng(x)} \gamma_T(dx)$. Similarly, the optimizer for the distributed control problem (2.11) is the Brownian bridge with terminal law Q^* .

Remark 2.16. Proposition 1.6 admits a similar control-theoretic formulation in terms of *deterministic* controls. Let V_{det} denote the value of the stochastic control problem (2.7) but with the supremum limited to those admissible pairs (α, X) in which the control is non-random, that is, $\alpha_i(t, x) = \tilde{\alpha}_i(t)$ for some $\tilde{\alpha}_i \in L^2[0, T]$. For these controls, X_t is Gaussian with covariance matrix tl for each $t \in [0, T]$. It can then be shown that

$$V_{\text{det}} = \sup_{y \in \mathbb{R}^n} \left(\int_{\mathbb{R}^n} g d\gamma_{y, T} - \frac{1}{n} H(\gamma_{y, T} | \gamma_T) \right) = \sup_{y \in \mathbb{R}^n} \left(\int_{\mathbb{R}^n} g d\gamma_{y, T} - \frac{|y|^2}{2nT} \right),$$

and Proposition 1.6 yields the following analogue of (2.13):

$$0 \leq V_{\text{orig}} - V_{\text{det}} \leq \frac{nT^2}{2} \sum_{i,j=1}^n \int_{\mathbb{R}^n} |\partial_{ij} g|^2 d\gamma_{y^*, T},$$

where $y^* \in \mathbb{R}^n$ is the unique solution of $y^* = T \int_{\mathbb{R}^n} \nabla g d\gamma_{y^*, T}$.

Remark 2.17. Subsequent to the first version of this paper, a followup work [43] by the first author and Jackson partially extended Corollary 2.14 to cover more general objective functions, of the

form

$$\mathbb{E} \left[g_1(X_T) - \int_0^T \left(g_2(X_t) + \frac{1}{n} \sum_{i=1}^n L^i(X_t^i, \alpha^i(t, X_t)) \right) dt \right].$$

The “static” representations of the value functions V_{orig} and V_{dstr} given in (2.10) and (2.12) are no longer available for non-quadratic running costs. For this reason, completely different techniques were developed in [43], relying on partial differential equations on the space of probability measures and a form of the stochastic maximum principle. The results of [43] only partially extend our Corollary 2.14, because when specialized to the quadratic cost the former require stronger regularity and growth assumptions.

3 Proof of the Main Theorem

The proofs will make use of the well known log-Sobolev and Poincaré inequalities for strongly log-concave measures, recalled here for convenience as we will use them in several parts of the paper. The former is due to Bakry-Émery (see [6] or [7, Corollary 5.7.2]), and the latter is a consequence of the Brascamp-Lieb inequality [19, Theorem 4.1].

Theorem 3.1 (Log-Sobolev inequality). If $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is C^2 and κ -concave, and $R(dx) = e^{h(x)} dx$ is a probability measure, then R satisfies the log-Sobolev inequality,

$$H(Q|R) \leq \frac{1}{2\kappa} \int_{\mathbb{R}^n} \left| \nabla \log \frac{dQ}{dR} \right|^2 dQ,$$

for every $Q \in \mathcal{P}(\mathbb{R}^n)$ such that $Q \ll R$ and the weak gradient of $\log dQ/dR$ exists in $L^2(Q)$.

Theorem 3.2 (Poincaré inequality). If $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is κ -concave, and $R(dx) = e^{h(x)} dx$ is a probability measure, then R satisfies the Poincaré inequality,

$$\text{Var}_R(\varphi) := \int_{\mathbb{R}^n} \varphi^2 dR - \left(\int_{\mathbb{R}^n} \varphi dR \right)^2 \leq \frac{1}{\kappa} \int_{\mathbb{R}^n} |\nabla \varphi|^2 dR,$$

for every continuously differentiable function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ in $L^1(R)$.

The above Poincaré inequality is normally stated with the additional assumptions that h is C^2 , which is easily removed by mollification by a Gaussian, and that $\varphi \in L^2(R)$, which can be weakened to $L^1(R)$ by monotone approximation, though both sides may be infinite.

We will also make use of the Gibbs variational principle, which is well known, but we give the proof as we need a non-standard form which is careful about edge cases. Recall our convention that $H(Q) := \infty$ if Q is not absolutely continuous or if $Q \log Q \notin L^1(\mathbb{R}^n)$.

Theorem 3.3 (Gibbs variational principle). Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ be measurable, bounded from above and such that $Z := \int_{\mathbb{R}^n} e^f dx \in (0, \infty)$. Define $P \in \mathcal{P}(\mathbb{R}^n)$ by $P(dx) = Z^{-1} e^{f(x)} dx$. Then

$$\sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right) = \log Z \in (-\infty, \infty), \quad (3.1)$$

and the following are equivalent:

- (1) $H(P) < \infty$.
- (2) The supremum in (3.1) is attained uniquely by P .
- (3) There exists a maximizer in (3.1).

Proof. We first prove (3.1). Since f is bounded from above, $\int_{\mathbb{R}^n} f dQ \in [-\infty, \infty)$ is well-defined for all $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$. We may thus restrict the supremum in (3.1) to those Q with $H(Q) < \infty$. For $H(Q) < \infty$, we have the simple identity

$$\int_{\mathbb{R}^n} f dQ - H(Q) = -H(Q|P) + \log Z. \quad (3.2)$$

Therefore,

$$\sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right) = -\inf \left\{ H(Q|P) : Q \in \mathcal{P}(\mathbb{R}^n), H(Q) < \infty \right\} + \log Z,$$

and it suffices to show that the infimum on the right-hand side is zero. We proceed by approximation. For each $k \in \mathbb{N}$, let $B_k \subset \mathbb{R}^n$ denote the centered ball of radius k , and define the probability density $Q_k = P1_{B_k}/P(B_k)$. Since f is bounded from above, the density Q_k is bounded and supported on the bounded set B_k . Thus $Q_k \log Q_k \in L^1(\mathbb{R}^n)$, or $H(Q_k) < \infty$, and we conclude that $H(Q|P) \leq \liminf_k H(Q_k|P)$. Finally, since $P(B_k) \rightarrow 1$,

$$H(Q_k|P) = -\log P(B_k) \rightarrow 0.$$

This proves the claim (3.1).

Turning to the equivalence of (1)–(3), the implication (1) $\Rightarrow$ (2) follows by taking $Q = P$ in (3.2). The implication (2) $\Rightarrow$ (3) is trivial. Lastly, for the implication (3) $\Rightarrow$ (1), suppose $Q \in \mathcal{P}(\mathbb{R}^n)$ attains the supremum in (3.1). We know from (3.1) that the supremum is not $-\infty$, so $H(Q) < \infty$. Then, for any $R \in \mathcal{P}(\mathbb{R}^n)$ with $H(R) < \infty$, the identity (3.2) implies

$$-H(R|P) + \log Z = \int_{\mathbb{R}^n} f dR - H(R) \leq \int_{\mathbb{R}^n} f dQ - H(Q) = -H(Q|P) + \log Z.$$

Rearrange and minimize over R to get

$$H(Q|P) \leq \inf \left\{ H(R|P) : R \in \mathcal{P}(\mathbb{R}^n), H(R) < \infty \right\} = 0,$$

where the last equality was shown just above while proving (3.1). It follows that $H(Q|P) = 0$, so $Q = P$, and $H(P) = H(Q) < \infty$. This completes the proof. ■

3.1 Proof of Theorem 1.1

This section proves Theorem 1.1 in several parts, and we assume throughout that f satisfies the assumptions therein. Since f is C^2 and κ -concave,

$$f(x) \leq a - b|x|^2, \quad \text{for all } x \in \mathbb{R}^n, \quad \text{where } a := f(0) + \kappa^{-1}|\nabla f(0)|^2, \quad b := \kappa/4. \quad (3.3)$$

This implies that $Z := \int_{\mathbb{R}^n} e^{f(x)} dx < \infty$, so $P(dx) = Z^{-1}e^{f(x)}dx$ is well defined. Moreover, f is bounded from above, so $\int_{\mathbb{R}^n} f dQ$ is well defined in $[-\infty, \infty)$ for every $Q \in \mathcal{P}(\mathbb{R}^n)$. Note lastly that $fe^f \in L^1(\mathbb{R}^n)$, or equivalently $H(P) < \infty$, which follows from the growth assumption on $|f|$ and the fact that the κ -log-concave measure P satisfies $\int_{\mathbb{R}^n} e^{c|x|^2} P(dx) < \infty$ for each $c < \kappa/2$. (In fact, every absolutely continuous log-concave measure has finite entropy [12, Theorem I.1].) We first establish some properties of the optimization and fixed point problems appearing in Theorem 1.1.

Lemma 3.4. It holds that

$$-\infty < \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right) < \infty, \quad (3.4)$$

and any $Q^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ attaining the supremum satisfies $f \in L^1(Q^*)$. Also, equation (1.12) is valid.

Proof. The Gibbs variational formula (Theorem 3.3) implies that the supremum in (3.4) is no greater than $\log Z < \infty$. To see that it is not $-\infty$, note that f is locally bounded because it is concave and real-valued. Hence, if Q is any product measure with bounded support and finite entropy (such as the uniform measure on $[0, 1]^n$), we can bound the supremum from below by $\int_{\mathbb{R}^n} f dQ - H(Q) > -\infty$. Now, if Q^* is an optimizer, then $H(Q^*) < \infty$ and $\int_{\mathbb{R}^n} f dQ^* > -\infty$, the latter implying that $f \in L^1(Q^*)$ since f is bounded from above.

To prove (1.12), note that the simple calculation (1.11) is valid for any $Q \in \mathcal{P}(\mathbb{R}^n)$ with $H(Q) < \infty$, though both sides are $+\infty$ if and only if $\int_{\mathbb{R}^n} f dQ = -\infty$. Since $\int_{\mathbb{R}^n} f dQ$ always exists in $[-\infty, \infty)$, the supremum in (3.4) remains the same when restricted to those Q with $H(Q) < \infty$. By infimizing (1.11) over $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ with finite entropy, we deduce that the left-hand side of (1.12) is finite and equals $\inf\{H(Q|P) : Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n), H(Q) < \infty\}$. To complete the proof, we claim that if $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ satisfies $H(Q|P) < \infty$ and $H(Q) = \infty$, then there exists $Q_k \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ such that $H(Q_k) < \infty$ for each k and $H(Q_k|P) \rightarrow H(Q|P)$. Indeed, define the probability density $Q_k = Q1_{C_k}/Q(C_k)$, where $C_k = [-k, k]^n$, for k large enough that $Q(C_k) > 0$. Then

$$H(Q_k|P) = \frac{1}{Q(C_k)} \int_{C_k} \log \frac{dQ}{dP} dQ - \log Q(C_k)$$

is finite and converges to $H(Q|P)$ as $k \rightarrow \infty$. In particular, $\log(dQ_k/dP) \in L^1(Q_k)$. We also have $\log P = f - \log Z \in L^1(Q_k)$ because f is locally bounded and Q_k has compact support. We deduce that $\log Q_k \in L^1(Q_k)$, or $H(Q_k) < \infty$, which completes the proof. ■

The following proposition shows essentially that the fixed point problem (1.3) is the first order condition for optimality in (1.4). This extends naturally to much more general settings, with $(\mathbb{R}^n, dx)$ replaced by a general σ -finite product measure space, but we will not need this.

Proposition 3.5 (Optimality to fixed point). Suppose $Q^* = Q_1^* \times \cdots \times Q_n^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ attains the supremum in (3.4). Then $f \in L^1(Q^*)$ and Q^* satisfies the fixed point equation

$$\begin{aligned} Q_i^*(dx_i) &= Z_i^{-1} e^{\hat{f}_i(x_i)} dx_i, \quad \text{where } \hat{f}_i : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\} \text{ is defined by} \\ \hat{f}_i(x_i) &:= \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) \prod_{j \neq i} Q_j^*(dx_j), \quad i \in [n]. \end{aligned} \quad (3.5)$$

Proof. Note that $f \in L^1(Q^*)$ by Lemma 3.4. By assumption, $(Q_1^*, \dots, Q_n^*)$ attains the supremum

$$\sup_{Q_1, \dots, Q_n \in \mathcal{P}(\mathbb{R})} \left(\int_{\mathbb{R}^n} f d(Q_1 \times \cdots \times Q_n) - H(Q_1 \times \cdots \times Q_n) \right).$$

Clearly, $\hat{f}_i(x_i) = \mathbb{E}_{Q^*}[f(X) | X_i = x_i]$ for Q_i^* -a.e. $x_i \in \mathbb{R}$. Also, it is well known that entropy tensorizes for product measures: $H(Q_1 \times \cdots \times Q_n) = \sum_{i=1}^n H(Q_i)$. From these and the tower property it follows for each $i \in [n]$ that Q_i^* attains the supremum

$$S_i := \sup_{Q_i \in \mathcal{P}(\mathbb{R})} \left(\int_{\mathbb{R}} \hat{f}_i dQ_i - H(Q_i) \right). \quad (3.6)$$

We wish to invoke the Gibbs variational principle (Theorem 3.3) to deduce that this supremum is uniquely attained by the probability measure with density proportional to $e^{\hat{f}_i}$, and thus $Q_i^*(dx_i) = Z_i^{-1} e^{\hat{f}_i(x_i)} dx_i$, which yields (3.5). It remains to carefully check the conditions of Theorem 3.3. We know that Q_i^* attains the supremum (3.6), so we must just check that $Z_i \in (0, \infty)$. Note that (3.3) implies $f(x) \leq a - bx_i^2$ for all $x \in \mathbb{R}^n$, and thus $\hat{f}_i(x_i) \leq a - bx_i^2$ for all $x_i \in \mathbb{R}$, which implies $Z_i = \int_{\mathbb{R}} e^{\hat{f}_i(x_i)} dx_i < \infty$. Next, recall from Lemma 3.4 that $f \in L^1(Q^*)$, so by Fubini's theorem, $Q_i^*(|\hat{f}_i| < \infty) = 1$. Note that Q_i^* is absolutely continuous since $H(Q_i^*) < \infty$. Hence, $\{|\hat{f}_i| < \infty\}$ has nonzero Lebesgue measure, and so $Z_i > 0$. ■

The following constructions will be useful in the proofs of Lemmas 3.6 and 3.9. Define $G_1, G_2, G : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}$ by

$$G_1(Q) := \int_{\mathbb{R}^n} f dQ, \quad G_2(Q) := H(Q), \quad G := G_1 - G_2. \quad (3.7)$$

For $R = R_1 \times \cdots \times R_n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ and $S = S_1 \times \cdots \times S_n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ with strictly positive density a.e., we denote by $M^{R \rightarrow S}(t) = (M_1^{R \rightarrow S}(t), \dots, M_n^{R \rightarrow S}(t))$ the displacement interpolations between the marginals, that is,

$$M_i^{R \rightarrow S}(t) = ((1-t)\text{Id} + tT_i^{R \rightarrow S})_{\#} R_i, \quad (3.8)$$

where $T_i^{R \rightarrow S} : \mathbb{R} \rightarrow \mathbb{R}$ is the R_i -a.s. unique nondecreasing function satisfying $S_i = (T_i^{R \rightarrow S})_{\#} R_i$.

Lemma 3.6. There exists a unique maximizer in (3.4).

Proof. We first prove existence. Recalling the identity (1.12), the optimizers of (3.4) are in one-to-one correspondence with the optimizers of $\inf_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} H(Q|P)$. The latter exist because $\mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ is a weakly closed subset of $\mathcal{P}(\mathbb{R}^n)$ and because $H(\cdot|P)$ has weakly compact sub-level sets (see, e.g., [33, Lemma 1.4.3(c)]).

We next prove uniqueness. Let $Q^* = Q_1^* \times \cdots \times Q_n^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ denote any optimizer of (3.4). Recalling (3.7), we see that Q^* is a maximizer of G , and we will show it must be the only one. Let $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ be distinct from Q^* . Since Q_i^* and Q_i are distinct, there exists i such that $T_i^{Q^* \rightarrow Q}$ is different from the identity map on a set with strictly positive Q_i^* -measure. Writing out the expression of G_1 ,

$$G_1(M^{Q^* \rightarrow Q}(t)) = \int_{\mathbb{R}^n} f \left((1-t)x_1 + tT_1^{Q^* \rightarrow Q}(x_1), \dots, (1-t)x_n + tT_n^{Q^* \rightarrow Q}(x_n) \right) \prod_{i=1}^n Q_i^*(dx_i),$$

we see that $t \mapsto G_1(M^{Q^* \rightarrow Q}(t))$ is strictly concave because f is strictly concave and $Q \neq Q^*$. Tensorization of entropy yields $G_2(M^{Q^* \rightarrow Q}(t)) = \sum_{i=1}^n H(M_i(t))$, and it is well known that differential entropy is displacement convex [57, Theorem 5.15(i)]. That is, $t \mapsto H(M_i^{Q^* \rightarrow Q}(t))$ is convex for each i . We deduce that $t \mapsto G(M^{Q^* \rightarrow Q}(t))$ is strictly concave. This proves uniqueness: if Q were also an optimizer, then $G(M^{Q^* \rightarrow Q}(1)) = G(Q) = G(Q^*) = G(M^{Q^* \rightarrow Q}(0))$ would imply $G(M^{Q^* \rightarrow Q}(t)) > G(Q^*)$ for some $t \in (0, 1)$. ■

Remark 3.7. We do not expect uniqueness in Lemma 3.6 to hold under mere concavity of f . The challenge is that the differential entropy functional is displacement convex, but not strictly so.

In some of the following proofs, some shorthand notation will be useful. For $Q \in \mathcal{P}(\mathbb{R}^n)$, let us write Q_{-i} for the marginal of $(X_j)_{j \neq i}$ under Q . For $x \in \mathbb{R}^n$ let us write $x_{-i} = (x_j)_{j \neq i}$ and, with some abuse of notation, $f(x) = f(x_i, x_{-i})$.

Lemma 3.8. If $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ satisfies the fixed point equation (3.5), then Q is κ -log-concave.

Proof. By (3.5), the density of Q is proportional to $e^{\hat{F}}$, where $\hat{F}(x) = \sum_{i=1}^n \hat{f}_i(x_i)$ and $\hat{f}_i$ is given by

$$\hat{f}_i(x_i) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) Q_{-i}(x_{-i}) dx_{-i}.$$

By the κ -concavity of f , for any $y, z \in \mathbb{R}$ and $t \in [0, 1]$, we have

$$\begin{aligned} & \hat{f}_i(tz + (1-t)y) + \frac{\kappa}{2}(tz + (1-t)y)^2 \\ &= \int_{\mathbb{R}^{n-1}} \left[f(tz + (1-t)y, x_{-i}) + \frac{\kappa}{2}(tz + (1-t)y)^2 \right] Q_{-i}(x_{-i}) dx_{-i} \\ &\geq \int_{\mathbb{R}^{n-1}} \left[tf(z, x_{-i}) + t\frac{\kappa}{2}z^2 + (1-t)f(y, x_{-i}) + (1-t)\frac{\kappa}{2}y^2 \right] Q_{-i}(x_{-i}) dx_{-i} \\ &= t\hat{f}_i(z) + t\frac{\kappa}{2}z^2 + (1-t)\hat{f}_i(y) + (1-t)\frac{\kappa}{2}y^2. \end{aligned}$$

This shows that $\hat{f}_i$ is κ -concave, and thus so is $\hat{F}$. ■

The next proposition, in conjunction with Proposition 3.5, shows that the optimizers of (3.4) and the solutions of the fixed point problem (1.3) are exactly the same.

Proposition 3.9 (Fixed point to optimality). Let $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ satisfy $f \in L^1(Q)$ and the fixed point problem (3.5). Then Q has strictly positive density a.e. and is a maximizer of (3.4).

Proof. We first show that Q has strictly positive density a.e. Since $Q = Q_1 \times \cdots \times Q_n$ satisfies the fixed point equation (3.5), each Q_i has a density with exponent

$$\hat{f}_i(x_i) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) \prod_{j \neq i} Q_j(dx_j) \geq -c_1 e^{c_2 x_i^2} \prod_{j \neq i} \int_{\mathbb{R}} e^{c_2 x_j^2} Q_j(x_j) dx_j$$

for every $x_i \in \mathbb{R}$. From Lemma 3.8 we know that Q is κ -log-concave. Since $c_2 < \kappa/2$, we deduce that $\int_{\mathbb{R}} e^{c_2 x_j^2} Q_j(x_j) < \infty$. Thus $\hat{f}_i(x_i) > -\infty$ for all $x_i \in \mathbb{R}$.

Let Q^* be an optimizer of $\sup\{G(R) : R \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)\}$, which exists uniquely by Lemma 3.6. By Proposition 3.5, we have $f \in L^1(Q^*)$, and Q^* satisfies the fixed point equation (3.5). The argument given in the previous paragraph implies that Q^* has a strictly positive density a.e. To complete the proof, we must show that $G(Q) \geq G(Q^*)$.

Recalling (3.7)–(3.8), we see that $G(M^{Q \rightarrow Q^*}(t)) = g_1(t) - g_2(t)$, where

$$g_1(t) := \int_{\mathbb{R}^n} f((1-t)x_1 + tT_1^{Q \rightarrow Q^*}(x_1), \dots, (1-t)x_n + tT_n^{Q \rightarrow Q^*}(x_n)) \prod_{i=1}^n Q_i(dx_i),$$

$$g_2(t) := H(M_1^{Q \rightarrow Q^*}(t) \times \cdots \times M_n^{Q \rightarrow Q^*}(t)) = \sum_{i=1}^n H(M_i^{Q \rightarrow Q^*}(t)).$$

Let us write g^+ for the right-derivative of a real-valued function g , when it exists. Note that $T_i^{Q \rightarrow Q^*}$ is a.e. differentiable, as it is monotone. Using [57, Theorem 5.30], we may compute the right-derivatives at zero as

$$g_1^+(0) = \sum_{i=1}^n \int_{\mathbb{R}^n} \partial f(x) (T_i^{Q \rightarrow Q^*}(x_i) - x_i) Q(x) dx,$$

$$g_2^+(0) = - \sum_{i=1}^n \int_{\mathbb{R}} \left((T_i^{Q \rightarrow Q^*})'(x_i) - 1 \right) Q_i(x_i) dx_i.$$

We wish to rewrite both terms in more useful forms.

We first claim that

$$\int_{\mathbb{R}^n} \partial f(x) (T_i^{Q \rightarrow Q^*}(x_i) - x_i) Q(x) dx = \int_{\mathbb{R}} \hat{f}'_i(x_i) (T_i^{Q \rightarrow Q^*}(x_i) - x_i) Q_i(x_i) dx_i, \quad (3.9)$$

where $\hat{f}_i$ is defined as in (3.5). To see this, note that $\hat{f}_i(x_i) = \mathbb{E}_Q[f(x_i, X_{-i})]$ for all $x_i \in \mathbb{R}$, so

$$\hat{f}_i^+(x_i) = \lim_{h \downarrow 0} h^{-1} \mathbb{E}_Q[f(x_i + h, X_{-i}) - f(x_i, X_{-i})].$$

By the concavity of f , the difference quotient $[f(x_i + h, X_{-i}) - f(x_i, X_{-i})]/h$ increases as $h \downarrow 0$, and it is bounded from below for $0 < h \leq h_0$ by $[f(x_i + h_0, X_{-i}) - f(x_i, X_{-i})]/h_0$, which has finite Q -expectation for a.e. choice of $h_0 > 0$ by Fubini's theorem since $f \in L^1(Q)$. Hence, by monotone convergence,

$$\hat{f}_i^+(x_i) = \mathbb{E}_Q[\partial f(x_i, X_{-i})]. \quad (3.10)$$

Moreover, this quantity is finite and nonincreasing in x_i because $\hat{f}_i$ is a concave real-valued function. In addition, $\hat{f}'_i = \hat{f}_i^+$ a.e. since concave functions are a.e. differentiable. Using (3.10), we see that the right-hand side of (3.9) equals $\mathbb{E}_Q[\mathbb{E}_Q[\partial f(X) | X_i](T_i(X_i) - X_i)]$, which yields (3.9).

We next integrate by parts to get

$$-\int_{\mathbb{R}} \left((T_i^{Q \rightarrow Q^*})'(x_i) - 1 \right) Q_i(x_i) dx_i = \int_{\mathbb{R}} \left(T_i^{Q \rightarrow Q^*}(x_i) - x_i \right) Q_i'(x_i) dx_i. \quad (3.11)$$

To justify this carefully, we use the following version of Lebesgue-Stieltjes integration by parts (see, e.g. [55, Proposition 4.5]): Let $-\infty < a < b < \infty$. For two functions U and V of finite variation, which are right-continuous with left-limits,

$$U(b)V(b) - U(a)V(a) = \int_a^b U(s-) dV(s) + \int_a^b V(s-) dU(s) + \sum_{u \in (a,b]} \Delta U_u \Delta V_u, \quad (3.12)$$

where $\Delta U_t = U(t) - U(t-)$. We will apply this to $U(x_i) = T_i(x_i) - x_i$ and $V(x_i) = Q_i(x_i)$. Note that the probability density function of Q_i is absolutely continuous on $[a, b]$ because it is proportional to $e^{\hat{f}_i}$, and $\hat{f}_i$ is Lipschitz on $[a, b]$ as a concave function. Let F_{Q_i} and $F_{Q_i^*}$ denote the CDFs of Q_i and Q_i^* respectively. Recalling that $T_i^{Q \rightarrow Q^*} = F_{Q_i^*}^{-1} \circ F_{Q_i}$ is the monotone map pushing Q_i forward to Q_i^* , and that both Q_i and Q_i^* admit strictly positive densities, we see that the function T_i is absolutely continuous. Hence, there is no jump term in the integration by parts (3.12), and we have

$$\begin{aligned} & (T_i^{Q \rightarrow Q^*}(b) - b)Q_i(b) - (T_i^{Q \rightarrow Q^*}(a) - a)Q_i(a) \\ &= \int_a^b \left(T_i^{Q \rightarrow Q^*}(x_i) - x_i \right) Q_i'(x_i) dx_i + \int_a^b \left((T_i^{Q \rightarrow Q^*}(x_i) - 1) Q_i(x_i) \right) dx_i. \end{aligned}$$

Letting $a \rightarrow -\infty$ and $b \rightarrow \infty$, we arrive at (3.11) as long as the boundary terms vanish. For this it suffices to show that there exist sequences $x_k^\pm \rightarrow \pm\infty$ such that

$$\lim_{k \rightarrow \infty} (T_i^{Q \rightarrow Q^*}(x_k^\pm) - x_k^\pm) Q_i(x_k^\pm) = 0.$$

If this were not the case, it would imply that $|T_i^{Q \rightarrow Q^*}(x) - x|Q_i(x) \leq (|T_i^{Q \rightarrow Q^*}(x)| + |x|)Q_i(x)$ is bounded away from zero for $|x|$ sufficiently large. This would in turn imply that $\int_{\mathbb{R}} (|T_i^{Q \rightarrow Q^*}(x_i)| + |x_i|)Q_i(x_i) dx_i = \infty$, contradicting the fact that

$$\int_{\mathbb{R}} (|T_i^{Q \rightarrow Q^*}(x_i)| + |x_i|)Q_i(x_i) dx_i = \int_{\mathbb{R}} |x_i|Q_i^*(x_i) dx_i + \int_{\mathbb{R}} |x_i|Q_i(x_i) dx_i < \infty.$$

Both integrals are finite because Q_i and Q_i^* are κ -log-concave by Lemma 3.8 and thus admit finite moments of every order. With (3.11) and (3.9) now justified, we see that the right-derivative of $G(M^{Q \rightarrow Q^*}(t))$ at $t = 0$ is

$$g_1^{'+}(0) - g_2^{'+}(0) = \sum_{i=1}^n \int_{\mathbb{R}} (\hat{f}_i'(x_i)Q_i(x_i) - Q_i'(x_i)) (T_i^{Q \rightarrow Q^*}(x_i) - x_i) dx_i.$$

This is in fact zero, because Q_i is proportional to $e^{\hat{f}_i}$. We saw in the proof of Lemma 3.6 that $G(M^{Q \rightarrow Q^*}(t))$ is concave. Since we now know that it has vanishing right-derivative at $t = 0$, it follows that $G(M^{Q \rightarrow Q^*}(1)) \leq G(M^{Q \rightarrow Q^*}(0))$. That is, $G(Q^*) \leq G(Q)$, which completes the proof. ■

Proof of Theorem 1.1. Let S_{opt} denote the set of maximizers in (3.4), and let S_{fix} denote the set of $Q^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ satisfying $f \in L^1(Q^*)$ and the fixed point equation (3.5). Proposition 3.5 shows that $S_{\text{opt}} \subset S_{\text{fix}}$. Proposition 3.9 shows conversely that $S_{\text{opt}} \supset S_{\text{fix}}$, so in fact $S_{\text{opt}} = S_{\text{fix}}$. Lemma 3.6 shows that this set is a singleton. Its unique element Q^* is κ -log-concave by Lemma 3.8 and has strictly positive density a.e. by Proposition 3.9. This proves claims (1–3) of Theorem 1.1.

To prove (4), recall the identity (1.12), which shows that

$$R_f = \log \int_{\mathbb{R}^n} e^{f(x)} dx - \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} f dQ - H(Q) \right) = H(Q^* | P).$$

The κ -log-concavity of P and the log-Sobolev inequality (Theorem 3.1) imply

$$H(Q^* | P) \leq \frac{1}{2\kappa} \int_{\mathbb{R}^n} \left| \nabla \log \frac{dQ^*}{dP} \right|^2 dQ^*.$$

Since $Q^* = Q_1^* \times \cdots \times Q_n^*$ is a product measure, we have $\partial_i \log Q^*(x) = \partial_i \log Q_i^*(x_i)$ for $x \in \mathbb{R}^n$ and note that the derivative exists almost everywhere because $\log Q_i^*$ is concave. We saw in (3.10) in the proof of Proposition 3.9 that the following identity is valid for almost every $x_i \in \mathbb{R}$, with the expectation on the right-hand side being finite:

$$\partial_i \log Q_i^*(x_i) = \partial_i \mathbb{E}_{Q^*}[f(X) | X_i = x_i] = \mathbb{E}_{Q^*}[\partial_i f(X) | X_i = x_i].$$

Thus,

$$\begin{aligned} H(Q^* | P) &\leq \frac{1}{2\kappa} \int_{\mathbb{R}^n} \sum_{i=1}^n \left| \partial_i \log Q_i^*(x_i) - \partial_i f(x) \right|^2 Q^*(dx) \\ &= \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n (\mathbb{E}_{Q^*}[\partial_i f(X) | X_i] - \partial_i f(X))^2 \\ &= \frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*}(\partial_i f(X) | X_i). \end{aligned}$$

This yields the first bound in (1.5). Recall that Q_{-i}^* denotes the law of $(X_j)_{j \neq i}$, which equals the conditional law of $(X_j)_{j \neq i}$ given X_i under Q^* by independence. The measure Q_{-i}^* is κ -log-concave because Q_j^* is for each j . Hence, it obeys a Poincaré inequality (Theorem 3.2), $\text{Var}_{Q_{-i}^*}(\varphi) \leq \kappa^{-1} \int_{\mathbb{R}^{n-1}} |\nabla \varphi|^2 dQ_{-i}^*$, for any C^1 function $\varphi \in L^1(Q_{-i}^*)$. Applying this to $\partial_i f$ with coordinate i fixed,

$$\text{Var}_{Q^*}(\partial_i f(X) | X_i) \leq \frac{1}{\kappa} \sum_{j \neq i} \mathbb{E}_{Q^*}[|\partial_{ij} f(X)|^2 | X_i].$$

Complete the proof of the second inequality of (1.5) by using the tower property to get

$$\frac{1}{2\kappa} \mathbb{E}_{Q^*} \sum_{i=1}^n \text{Var}_{Q^*}(\partial_i f(X) | X_i) \leq \frac{1}{2\kappa^2} \sum_{i=1}^n \sum_{j \neq i} \mathbb{E}_{Q^*}[|\partial_{ij} f(X)|^2].$$

3.2 Proof of Corollary 1.4

Let $f(x) := g(x) + \sum_{i=1}^n V_i(x_i)$. Then $\int_{\mathbb{R}^n} e^g d\rho = \int_{\mathbb{R}^n} e^{f(x)} dx$, and the concavity of g and κ -concavity of V_i imply that f is κ -concave. Note also that for any $Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} g dQ - H(Q | \rho) = \int_{\mathbb{R}^n} f dQ - H(Q).$$

This shows that the optimization problems (1.4) and (1.9) are the same. Moreover, the fixed point problems (1.8) and (1.3) admit exactly the same solutions: Q_i^* solves (1.3) if and only if it solves (1.8). With these identifications, applying Theorem 1.1 to f immediately proves claims (1–3) of Corollary 1.4.

Finally, with Q_i^* solving (3.5) (or equivalently (1.8)), we have

$$R_g^p = R_f \leq \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q_i^*} [|\partial_{ij} f(X)|^2] = \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q_i^*} [|\partial_{ij} g(X)|^2],$$

because $\partial_{ij} f = \partial_{ij} g$ for all $i \neq j$. This proves claim (4) of Corollary 1.4.

3.3 Proof of Proposition 1.6

Note that $\int_{\mathbb{R}^n} f(x+y) \gamma_t(dx) < \infty$ for each $y \in \mathbb{R}^n$ by the growth assumption on f . The function

$$y \mapsto \int_{\mathbb{R}^n} f d\gamma_{y,t} - H(\gamma_{y,t} | \gamma_t) = \int_{\mathbb{R}^n} f(x+y) \gamma_t(dx) - \frac{1}{2t} |y|^2$$

is $(1/t)$ -concave and thus bounded from above. It admits a unique maximizer obtained by setting the gradient equal to zero; the first order condition is precisely (1.18). Let $P(dx) = Z^{-1} e^{f(x)} \gamma_t(dx)$. The simple identity

$$\log \int_{\mathbb{R}^n} e^f d\gamma_t - \left(\int_{\mathbb{R}^n} f d\gamma_{y,t} - H(\gamma_{y,t} | \gamma_t) \right) = H(\gamma_{y,t} | P),$$

valid for all $y \in \mathbb{R}^n$, implies that

$$\log \int_{\mathbb{R}^n} e^f d\gamma_t - \sup_{y \in \mathbb{R}^n} \left(\int_{\mathbb{R}^n} f d\gamma_{y,t} - H(\gamma_{y,t} | \gamma_t) \right) = \inf_{y \in \mathbb{R}^n} H(\gamma_{y,t} | P).$$

The right-hand side is equal to $H(\gamma_{y^*,t} | P)$. The measure P is $(1/t)$ -log-concave, so we may use the log-Sobolev inequality (Theorem 3.1) to get

$$\begin{aligned} H(\gamma_{y^*,t} | P) &\leq \frac{t}{2} \int_{\mathbb{R}^n} \left| \nabla \log \frac{d\gamma_{y^*,t}}{dP} \right|^2 d\gamma_{y^*,t} = \frac{t}{2} \int_{\mathbb{R}^n} \left| \nabla \log \frac{d\gamma_{y^*,t}}{d\gamma_t} - \nabla \log \frac{dP}{d\gamma_t} \right|^2 d\gamma_{y^*,t} \\ &= \frac{t}{2} \int_{\mathbb{R}^n} \left| \frac{1}{t} y^* - \nabla f(x) \right|^2 \gamma_{y^*,t}(dx) \\ &= \frac{t}{2} \sum_{i=1}^n \text{Var}_{\gamma_{y^*,t}}(\partial_i f), \end{aligned}$$

where the last step follows from (1.18). Using the Gaussian Poincaré inequality (or Theorem 3.2), this is bounded by the second term on the right-hand side of (1.19).

3.4 Asymptotic independence

Proof of first inequality in (1.16). Let $P, Q \in \mathcal{P}(\mathbb{R}^n)$. Let $k_1, \dots, k_m$ be positive integers summing to n . Suppose $P_1, \dots, P_m$ are the marginals of P on $\mathbb{R}^{k_1}, \dots, \mathbb{R}^{k_m}$, and define the marginals $Q_1, \dots, Q_m$ similarly. Then

$$\sum_{i=1}^m \mathcal{W}_2^2(P_i, Q_i) \leq \mathcal{W}_2^2(P, Q).$$

Indeed, to prove this, let (X, Y) be an optimal coupling of (P, Q) . Let X_i be the $\mathbb{R}^{k_i}$ coordinate, for $i = 1, \dots, m$, and similarly define Y_i . Then (X_i, Y_i) is a coupling of (P_i, Q_i) , and so

$$\mathcal{W}_2^2(P, Q) = \mathbb{E}[|X - Y|^2] = \mathbb{E}\left[\sum_{i=1}^m |X_i - Y_i|^2\right] \geq \sum_{i=1}^m \mathcal{W}_2^2(P_i, Q_i).$$

Now, let $1 \leq k \leq n$, and let $m = \lfloor n/k \rfloor$. Let Π be the set of vectors $(S_1, \dots, S_m)$ of disjoint k -element subsets of $[n]$. Let Q_{S_i} and P_{S_i} denote the corresponding marginals, on those coordinates in $S_i \subset [n]$. Note

that $\mathcal{W}_2^2(P_{S_i}, Q_{S_i})$ does not depend on the order of the elements of S_i . Then

$$\sum_{i=1}^m \mathcal{W}_2^2(P_{S_i}, Q_{S_i}) \leq \mathcal{W}_2^2(P_{S_1 \cup \dots \cup S_m}, Q_{S_1 \cup \dots \cup S_m}) \leq \mathcal{W}_2^2(P, Q).$$

If $(S_1, \dots, S_m)$ is chosen uniformly at random from Π and i is chosen uniformly at random from $[m]$, then the marginal law of S_i is the same as the law of a uniformly random choice of k -element subset of $[n]$. In particular,

$$\frac{1}{\binom{n}{k}} \sum_{S \subset [n], |S|=k} \mathcal{W}_2^2(P_S, Q_S) = \frac{1}{|\Pi|} \sum_{(S_1, \dots, S_m) \in \Pi} \frac{1}{m} \sum_{i=1}^m \mathcal{W}_2^2(P_{S_i}, Q_{S_i}).$$

Combining the two previous inequalities yields

$$\frac{1}{\binom{n}{k}} \sum_{S \subset [n], |S|=k} \mathcal{W}_2^2(P_S, Q_S) \leq \frac{1}{m} \mathcal{W}_2^2(P, Q) = \frac{1}{[n/k]} \mathcal{W}_2^2(P, Q).$$

Proof of Corollary 1.2. By the triangle inequality, the square root of the left-hand side of (1.7) is no more than $A_1 + A_2$, where we define

$$A_1 := \mathbb{E}_P \left[\left(\frac{1}{n} \sum_{i=1}^n \varphi(X_i) - \frac{1}{n} \sum_{i=1}^n \mathbb{E}_P[\varphi(X_i)] \right)^2 \right]^{1/2},$$

$$A_2 := \left| \frac{1}{n} \sum_{i=1}^n (\mathbb{E}_P[\varphi(X_i)] - \mathbb{E}_{Q^*}[\varphi(X_i)]) \right|.$$

Recall that $|\varphi'| \leq 1$. Using Kantorovich duality and (1.16) with $k = 1$,

$$A_2^2 \leq \frac{1}{n} \sum_{i=1}^n \mathcal{W}_1^2(P_i, Q_i^*) \leq \frac{1}{n} \sum_{i=1}^n \mathcal{W}_2^2(P_i, Q_i^*) \leq \frac{2R_f}{\kappa n}.$$

Apply the Poincaré inequality (Theorem 3.2) to the function $x \mapsto (1/n) \sum_{i=1}^n \varphi(x_i)$ to get

$$A_1^2 = \text{Var}_P \left(\frac{1}{n} \sum_{i=1}^n \varphi(X_i) \right) \leq \frac{1}{\kappa n^2} \sum_{i=1}^n \mathbb{E}_P[|\varphi'(X_i)|^2] \leq \frac{1}{\kappa n}.$$

Combine these two bounds to complete the proof.

4 Gibbs Measure Proofs

This section proves the results of Section 2.1. Throughout, the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined as in (1.6) and satisfies Assumption 2.1.

Proof of Lemma 2.2. Compute two derivatives to find, for all $i \neq j$,

$$\partial_{ij} f(x) = V''(x_i) + \sum_{j \neq i} J_{ij} K''(x_i - x_j), \quad \partial_{ij} f(x) = -J_{ij} K''(x_i - x_j).$$

Hence, for any $x, z \in \mathbb{R}^n$,

$$z^\top \nabla^2 f(x) z = \sum_{i,j=1}^n z_i z_j \partial_{ij} f(x) = \sum_{i=1}^n z_i^2 V''(x_i) + \sum_{i,j=1}^n (z_i^2 - z_i z_j) J_{ij} K''(x_i - x_j).$$

Using the evenness of K'' and the symmetry of J ,

$$\sum_{i,j=1}^n (z_i^2 - z_i z_j) J_{ij} K''(x_i - x_j) = \frac{1}{2} \sum_{i,j=1}^n (z_i - z_j)^2 J_{ij} K''(x_i - x_j).$$

Since $K'' \leq 0$ and $J_{ij} \geq 0$, we find that this quantity is nonpositive. By κ -concavity of V ,

$$z^\top \nabla^2 f(x) z \leq \sum_{i=1}^n z_i^2 V''(x_i) \leq -\kappa |z|^2,$$

which shows that f is κ -concave.

Proof of Corollary 2.3. Note that f is C^2 and κ -concave. Also, the assumptions on $|V|$ and $|K'|$ in Assumption 2.1 clearly imply that $|f|$ satisfies the growth assumption in Theorem 1.1. Therefore, Theorem 1.1 applies. Let Q^* be given as therein. Computing derivatives as above, we have

$$R_f \leq \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} \mathbb{E}_{Q^*} [|\partial_{ij} f(X)|^2] = \frac{1}{\kappa^2} \sum_{1 \leq i < j \leq n} J_{ij}^2 \mathbb{E}_{Q^*} [|K''(X_i - X_j)|^2]. \quad (4.1)$$

Using the assumption on K'' , we find

$$\mathbb{E}_{Q^*} [|K''(X_i - X_j)|^2] \leq a \mathbb{E}_{Q^*} [e^{b|X_i - X_j|}]. \quad (4.2)$$

By assumption, $X_i - X_j$ has mean zero under Q^* . It follows from the κ -log-concavity of Q^* that the law of $X_i - X_j$ is $(\kappa/2)$ -log-concave (see, e.g., [56, Theorem 3.7(a) and Theorem 3.8]). This implies that it is subgaussian in the sense that

$$\mathbb{E}_{Q^*} [e^{s(X_i - X_j)}] \leq e^{s^2/\kappa}, \quad \forall s \in \mathbb{R}.$$

Indeed, this can be deduced from the log-Sobolev inequality (Theorem 3.1) via Herbst's argument or [13, Theorem 1.3]. Thus, using (4.2),

$$\mathbb{E}_{Q^*} [|K''(X_i - X_j)|^2] \leq a \mathbb{E}_{Q^*} [e^{b(X_i - X_j)} + e^{b(X_j - X_i)}] \leq 2ae^{b^2/\kappa}.$$

Combine this with (4.1) to complete the proof.

4.1 Doubly stochastic matrices

We now turn to the proof of Theorem 2.5. We first need a straightforward lemma about displacement convexity, which is a special case of [1, Propositions 7.5 and 7.7]. We include a direct proof here for the sake of completeness.

Lemma 4.1. Let $Q_1, \dots, Q_n \in \mathcal{P}(\mathbb{R})$ and $t_1, \dots, t_n \in [0, 1]$ be such that $\sum_{i=1}^n t_i = 1$. Then there exists a random vector $X = (X_1, \dots, X_n)$ such that $X_i \sim Q_i$ for each i and

$$H\left(\text{Law}\left(\sum_{i=1}^n t_i X_i\right)\right) \leq \sum_{i=1}^n t_i H(Q_i).$$

Proof. The proof is by induction on n , with the case $n = 1$ holding trivially. Assume that the statement of the lemma is true for some n . Let $Q_1, \dots, Q_{n+1} \in \mathcal{P}(\mathbb{R})$ and $t_1, \dots, t_{n+1} \in [0, 1]$ be such that $\sum_{i=1}^{n+1} t_i = 1$. Without loss of generality, assume that $t_{n+1} < 1$ and that $Q_1, \dots, Q_{n+1}$ have finite entropy, as otherwise

there is nothing to prove. For $i = 1, \dots, n$, define $\tilde{t}_i := t_i/(1 - t_{n+1})$, so that $\sum_{i=1}^n \tilde{t}_i = 1$. By assumption, we may find a random vector $(X_1, \dots, X_n)$ such that $X_i \sim Q_i$ for each $i = 1, \dots, n$ and

$$H(\tilde{Q}) \leq \sum_{i=1}^n \tilde{t}_i H(Q_i), \quad (4.3)$$

where $\tilde{Q}$ denotes the law of $\tilde{X} := \sum_{i=1}^n \tilde{t}_i X_i$. By absolute continuity, there is a unique nondecreasing function $T: \mathbb{R} \rightarrow \mathbb{R}$ such that $T_{\#}\tilde{Q} = Q_{n+1}$. The entropy functional is displacement convex [57, Theorem 5.15(i)], which means that the function

$$[0, 1] \ni t \mapsto H((tT + (1-t)\text{Id})_{\#}\tilde{Q})$$

is convex. In particular, letting $X_{n+1} = T(\tilde{X})$, we find

$$\begin{aligned} H(\text{Law}(t_{n+1}X_{n+1} + (1 - t_{n+1})\tilde{X})) &= H((t_{n+1}T + (1 - t_{n+1})\text{Id})_{\#}\tilde{Q}) \\ &\leq t_{n+1}H(Q_{n+1}) + (1 - t_{n+1})H(\tilde{Q}). \end{aligned}$$

By (4.3) and the definition of $\tilde{t}_i$, we have $(1 - t_{n+1})H(\tilde{Q}) \leq \sum_{i=1}^n t_i H(Q_i)$, completing the proof. $\blacksquare$

Proof of Theorem 2.5(1). Let us abbreviate

$$M_n := \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} M_n(Q), \quad (4.4)$$

where we define

$$\begin{aligned} M_n(Q) &:= \int_{\mathbb{R}^n} f dQ - H(Q) \\ &= \sum_{i=1}^n \int_{\mathbb{R}} V(x) Q_i(dx) + \frac{1}{2} \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q_i(dx) Q_j(dy) - \sum_{i=1}^n H(Q_i), \end{aligned}$$

where the last equality used the symmetry of J and K , the fact that the diagonal entries of J are zero, and the tensorization of entropy. Recall that $\log \int_{\mathbb{R}^n} e^f dx = M_n + R_f$, by definition of R_f . We will complete the proof by showing that

$$M_n = n \sup_{Q \in \mathcal{P}(\mathbb{R})} \left(\int_{\mathbb{R}} V dQ + \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q(dx) Q(dy) - H(Q) \right), \quad (4.5)$$

and that the optimizer $Q^* = Q_1^* \times \dots \times Q_n^*$ in (1.4) must be i.i.d. copies of the optimizer Q of (4.5), that is, $Q_1^* = \dots = Q_n^* = Q$. The latter claim proves part (3) of the theorem. We deduce part (1) from (4.5) by noting first that the i.i.d. form of Q^* implies $\mathbb{E}_{Q^*}[X_i - X_j] = 0$ for all i, j . Using this and the assumption $\text{Tr}(J^2) = o(n)$, we then apply Corollary 2.3 to deduce that $R_f/n \rightarrow 0$, and Theorem 2.5(1) follows.

The proof of the inequality $(\geq)$ in (4.5) is immediate upon restricting the supremum in (4.4) to i.i.d. measures and using $\sum_{i,j=1}^n J_{ij} = n$:

$$\begin{aligned} M_n &\geq \sup_{Q \in \mathcal{P}(\mathbb{R})} \left(n \int_{\mathbb{R}} V(x) Q(dx) + \frac{1}{2} \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q(dx) Q(dy) - nH(Q) \right) \\ &= n \sup_{Q \in \mathcal{P}(\mathbb{R})} \left(\int_{\mathbb{R}} V dQ + \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q(dx) Q(dy) - H(Q) \right). \end{aligned}$$

To prove the inequality ($\leq$) in (4.5), fix $Q = Q_1 \times \cdots \times Q_n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ arbitrarily. By Lemma 4.1, there exists a random vector $X = (X_1, \dots, X_n)$ such that $X_i \sim Q_i$ for all i and

$$H(\bar{Q}) \leq \frac{1}{n} \sum_{i=1}^n H(Q_i). \quad (4.6)$$

where $\bar{Q}$ denotes the law of $\frac{1}{n} \sum_{i=1}^n X_i$. Using the concavity of V , we find

$$\frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} V(x) Q_i(dx) = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n V(X_i) \right] \leq \int_{\mathbb{R}} V d\bar{Q}. \quad (4.7)$$

Let $Y = (Y_1, \dots, Y_n)$ be an independent copy of X . Using the concavity of K and the fact that $\sum_i J_{ij} = \sum_j J_{ij} = 1$, we have

$$\begin{aligned} & \frac{1}{n} \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) Q_i(dx) Q_j(dy) \\ &= \mathbb{E} \left[\frac{1}{n} \sum_{i,j=1}^n J_{ij} K(X_i - Y_j) \right] \leq \mathbb{E} \left[K \left(\frac{1}{n} \sum_{i,j=1}^n J_{ij} (X_i - Y_j) \right) \right] \\ &= \mathbb{E} \left[K \left(\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{j=1}^n Y_j \right) \right] = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \bar{Q}(dx) \bar{Q}(dy). \end{aligned} \quad (4.8)$$

Combining (4.6), (4.7), and (4.8), we see that

$$M_n(Q)/n \leq \int_{\mathbb{R}} V d\bar{Q} + \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} K(x-y) \bar{Q}(dx) \bar{Q}(dy) - H(\bar{Q}) = M_n(\bar{Q}^{\otimes n})/n.$$

In other words, for an arbitrary choice of product measure Q , we may increase $M_n(Q)$ by replacing Q with the i.i.d. measure $\bar{Q}^{\otimes n}$. This completes the proof.

Proof of Theorem 2.5(2). We first justify the uniqueness claim. From part (1.1) of Theorem 1.1, we know that the optimizer $Q^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ in (4.4) is unique. It follows from the previous paragraph that this unique optimizer is in fact i.i.d., that is, $Q^* = Q^{\otimes n}$, where $Q \in \mathcal{P}(\mathbb{R})$ is the (necessarily unique) optimizer of (4.5), which does not depend on n . This proves the desired uniqueness.

Turning to the proof of (2.2), recall that $R_f/n \rightarrow 0$, and use Corollary 1.2 and the aforementioned i.i.d. form of the optimizer $Q^* = Q^{\otimes n}$ to deduce that, for any 1-Lipschitz function φ ,

$$\mathbb{E}_P \left[\left(\frac{1}{n} \sum_{i=1}^n \varphi(X_i) - \int_{\mathbb{R}} \varphi dQ \right)^2 \right] \leq \frac{(1 + \sqrt{2R_f})^2}{\kappa n} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

This is enough to deduce that $\frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ converges to Q weakly in law.

4.2 Graphons proofs

This section is devoted to the proof of Theorem 2.8. For $W \in \mathcal{W}$ and any measurable function $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ bounded from above, define $T_{W,\psi} : \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R}) \rightarrow \mathbb{R}$ by

$$T_{W,\psi}(\mu) := \mathbb{E}_{\mu^{\otimes 2}} [\psi(X_1, X_2) W(U_1, U_2)].$$

where (U_1, X_1) and (U_2, X_2) are independent with law μ . Note that $W \geq 0$ is integrable, so $T_{W,\psi}(\mu)$ is well-defined in $[-\infty, \infty)$. Let $\bar{\mu} := \text{Unif}[0, 1] \times \rho$, and define $I : \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R}) \rightarrow [0, \infty]$ by

$$I(\mu) := H(\mu | \bar{\mu}) = \int_0^1 H(\mu_u | \rho) du,$$

with the second identity coming from the chain rule for relative entropy [33, Theorem B.2.1], and we recall that $\rho(dx) = e^{V(x)}dx$ is a probability measure. We begin with two lemmas pertaining to the continuity of $T_{W,\psi}$.

Lemma 4.2. Let $\mathcal{K} \subset \mathbb{R}$ be a compact interval. Let $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ be supported on $\mathcal{K}^2$ and continuous when restricted to $\mathcal{K}^2$.

(1) If $\{W_\ell\}$ converges to W in strong cut metric and $W_\ell, W \geq 0$, then

$$\sup_{\mu \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathbb{R})} |T_{W_\ell, \psi}(\mu) - T_{W, \psi}(\mu)| \rightarrow 0.$$

(2) The map $\mu \rightarrow T_{W, \psi}(\mu)$ is continuous on $\{\mu \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathbb{R}) : I(\mu) < \infty\}$, with respect to the topology of weak convergence.

Proof. We begin with (1). Let $\mathcal{V}$ denote the space of functions $\phi : \mathcal{K}^2 \mapsto \mathbb{R}$ of the form

$$\phi(x, y) = \sum_{i=1}^L c_i a_i(x) b_i(y), \quad (4.9)$$

for some $L \in \mathbb{N}$, $c_i \in \mathbb{R}$, and continuous functions $a_i, b_i : \mathcal{K} \rightarrow [0, 1]$. It is easy to check that $\mathcal{V}$ is closed under multiplication, contains the constant functions, separates points in $\mathcal{K}^2$, and is a vector subspace of the space $C(\mathcal{K}^2)$ of continuous real-valued functions on $\mathcal{K}^2$. By the Stone-Weierstrass Theorem, we deduce that $\mathcal{V}$ is dense in $C(\mathcal{K}^2)$ with the supremum norm. Let $\varepsilon > 0$, and find $\phi \in \mathcal{V}$ such that $|\psi - \phi| < \varepsilon$ uniformly on $\mathcal{K}^2$. Extend the domain of ϕ to $\mathbb{R}^2$ by setting $\phi = 0$ on the complement of $\mathcal{K}^2$. Then for all $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$,

$$|T_{W_\ell, \psi}(\mu) - T_{W_\ell, \phi}(\mu)| \leq \varepsilon \|W_\ell\|_{L_1[0,1]^2}, \quad |T_{W, \psi}(\mu) - T_{W, \phi}(\mu)| \leq \varepsilon \|W\|_{L_1[0,1]^2}.$$

Consequently, using the triangle inequality, we have

$$|T_{W_\ell, \psi}(\mu) - T_{W, \psi}(\mu)| \leq \varepsilon \|W_\ell\|_{L_1[0,1]^2} + \varepsilon \|W\|_{L_1[0,1]^2} + |T_{W_\ell, \phi}(\mu) - T_{W, \phi}(\mu)|. \quad (4.10)$$

Since ϕ is of the form (4.9), we have

$$T_{W_\ell, \phi}(\mu) = \sum_{i=1}^L c_i \int_{[0,1]^2} \bar{a}_i(u) \bar{b}_i(v) W_\ell(u, v) du dv,$$

where we define $\bar{a}_i(u) := \mathbb{E}_\mu[a_i(X) | U = u]$, and $\bar{b}_i$ similarly. This yields

$$|T_{W_\ell, \phi}(\mu) - T_{W, \phi}(\mu)| \leq \sum_{i=1}^L |c_i| d_\square(W_\ell, W). \quad (4.11)$$

Noting that $d_\square(W_\ell, W) \rightarrow 0$ implies $\|W_\ell\|_{L_1[0,1]^2} \rightarrow \|W\|_{L_1[0,1]^2}$, we may now combine (4.10) and (4.11), sending $\ell \rightarrow \infty$ and then $\varepsilon \rightarrow 0$, to prove the claim (1).

To prove (2), let μ_k be a sequence of measures in $\mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ converging weakly to μ_∞ , such that $I(\mu_\infty) < \infty$. Let W_ℓ be a sequence of continuous functions in $\mathcal{W}$ converging in $L_1[0, 1]^2$ to W . By the triangle inequality,

$$|T_{W, \psi}(\mu_k) - T_{W, \psi}(\mu_\infty)| \leq 2 \sup_{v \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathbb{R})} |T_{W, \psi}(v) - T_{W_\ell, \psi}(v)| + |T_{W_\ell, \psi}(\mu_k) - T_{W_\ell, \psi}(\mu_\infty)|.$$

The first term converges to 0 as $\ell \rightarrow \infty$, by part (1) and the fact that convergence in $L_1[0, 1]^2$ implies convergence in strong cut metric. The second term converges to 0 for fixed ℓ as $k \rightarrow \infty$, using the

fact that μ_k converges weakly to μ_∞ , and the set of discontinuity points of $W_\ell(\cdot, \cdot)\psi(\cdot, \cdot)$ is contained in $[0, 1]^2 \times \partial(\mathcal{K}^2)$, which has measure 0 under $\mu_\infty^{\otimes 2}$ (as μ_∞ is absolutely continuous with respect to Lebesgue measure on $[0, 1] \times \mathbb{R}$). ■

Lemma 4.3. Suppose μ_m is a sequence of measures in $\mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ converging weakly to μ_∞ . Let $\psi : \mathbb{R}^2 \mapsto \mathbb{R}$ be a continuous function, and let $W \in L_1[0, 1]^2$. For $1 \leq m \leq \infty$, let $(U_1^m, X_1^m), (U_2^m, X_2^m) \stackrel{\text{i.i.d.}}{\sim} \mu_m$. Then

$$W(U_1^m, U_2^m)\psi(X_1^m, X_2^m) \xrightarrow{d} W(U_1^\infty, U_2^\infty)\psi(X_1^\infty, X_2^\infty).$$

Proof. If W is continuous, then the claim is immediate. For a general W , we proceed as follows: Fix $\varepsilon > 0$, and let $\mathcal{K}$ be a compact set such that $\mathbb{P}(X_1^m \in \mathcal{K}, X_2^m \in \mathcal{K}) \geq 1 - \varepsilon$, which is again possible by tightness of $\{(X_1^m, X_2^m)\}_{m \in \mathbb{N}}$. Let g be a continuous function with

$$\|W - g\|_{L_1[0, 1]^2} < \frac{\varepsilon}{1 \vee \sup_{x, y \in \mathcal{K}} |\psi(x, y)|}.$$

Then on the event $\{X_1^m \in \mathcal{K}, X_2^m \in \mathcal{K}\}$, we have

$$|W(U_1^m, U_2^m)\psi(X_1^m, X_2^m) - g(U_1^m, U_2^m)\psi(X_1^m, X_2^m)| \leq \varepsilon.$$

Thus, for any continuous function $\phi : \mathbb{R} \rightarrow [0, 1]$ which is 1-Lipschitz, we have

$$\left| \mathbb{E}\phi(W(U_1^m, U_2^m)\psi(X_1^m, X_2^m)) - \mathbb{E}\phi(g(U_1^m, U_2^m)\psi(X_1^m, X_2^m)) \right| \leq 2\varepsilon.$$

Finally,

$$\mathbb{E}\phi(g(U_1^m, U_2^m)\psi(X_1^m, X_2^m)) \rightarrow \mathbb{E}\phi(g(U_1^\infty, U_2^\infty)\psi(X_1^\infty, X_2^\infty)),$$

by the result for continuous functions. Thus

$$\limsup_{m \rightarrow \infty} \left| \mathbb{E}\phi(W(U_1^m, U_2^m)\psi(X_1^m, X_2^m)) - \mathbb{E}\phi(W(U_1^\infty, U_2^\infty)\psi(X_1^\infty, X_2^\infty)) \right| \leq 4\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the proof of the lemma is complete. ■

Proof of Theorem 2.8(1).

We begin with some notation. For a measurable function $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ which is bounded from above, define $M_n^\psi := \sup_{Q \in \mathcal{P}_{\text{Pr}}(\mathbb{R}^n)} M_n^\psi(Q)$, where

$$\begin{aligned} M_n^\psi(Q) &:= \sum_{i=1}^n \int_{\mathbb{R}} V(x) Q_i(dx) + \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(x, y) Q_i(dx) Q_j(dy) - \sum_{i=1}^n H(Q_i) \\ &= \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(x, y) Q_i(dx) Q_j(dy) - \sum_{i=1}^n H(Q_i | \rho). \end{aligned} \quad (4.12)$$

Letting $\tilde{K} : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $\tilde{K}(x, y) = K(x - y)/2$, we are most interested in the choice $\psi = \tilde{K}$, but treating a general ψ will be helpful for a truncation argument. Let Q^* be as in Theorem 1.1. With this notation, we have $\log \int_{\mathbb{R}^n} e^{f(x)} dx = M_n^{\tilde{K}} + R_f$. Corollary 2.3 and the assumption that $\text{Tr}(J^2) = o(n)$ imply that $R_f/n \rightarrow 0$, and to prove Theorem 2.8 it will thus suffice to show that

$$\lim_{n \rightarrow \infty} M_n^\psi/n = \sup_{\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})} (T_{W, \psi}(\mu) - I(\mu)) \quad (4.13)$$

for any continuous function $\psi \leq 0$.

To this effect, use the assumption that $\{n\}_{n \geq 1}$ converges in weak cut metric to W to conclude the existence of a sequence of permutations $\{\pi_n\}_{n \geq 1}$ with $\pi_n \in S_n$, such that $\{nJ^{(\pi_n)}\}_{n \geq 1}$ converges in strong cut metric to W , where $J_{ij}^{(\pi_n)} := J_{\pi_n(i)\pi_n(j)}$ for $1 \leq i, j \leq n$. Since π_n is a permutation, for any $Q = Q_1 \times \cdots \times Q_n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ we can write

$$M_n^\psi(Q) = \sum_{i=1}^n \int_{\mathbb{R}} V(x) \tilde{Q}_i(dx) + \sum_{i,j=1}^n J_{\pi_n(i)\pi_n(j)} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(x, y) \tilde{Q}_i(dx) \tilde{Q}_j(dy) - \sum_{i=1}^n H(\tilde{Q}_i),$$

where $\tilde{Q}_i := Q_{\pi_n(i)} \in \mathcal{P}(\mathbb{R})$. Thus

$$\sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} M_n^\psi(Q) = \sup_{\tilde{Q} \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \tilde{M}_n^\psi(\tilde{Q}),$$

where $\tilde{M}_n^\psi(\cdot)$ defined similarly to $M_n^\psi(\cdot)$ in (4.12), but with J replaced by $J^{(\pi_n)}$. Since $nJ^{(\pi_n)}$ converges to W in strong cut metric, by replacing J with $J^{(\pi_n)}$ without loss of generality we assume throughout the rest of the proof that nJ converges in strong cut metric to W .

To prove (4.13), we need the following construction which essentially embeds $\mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ into $\mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ for all n . For any $Q = Q_1 \times \cdots \times Q_n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$, define a probability measure $\mu_n(Q) \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ as follows: If $(U, X) \sim \mu_n(Q)$, then $U \sim \text{Unif}[0, 1]$, and the conditional law of X given $\{(i-1)/n < U \leq i/n\}$ is given by Q_i . Then we have

$$\begin{aligned} T_{W_{nj}, \psi}(\mu_n(Q)) &= \frac{1}{n} \sum_{i,j=1}^n J_{ij} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(x, y) Q_i(dx) Q_j(dy), \\ I(\mu_n(Q)) &= \frac{1}{n} \sum_{i=1}^n H(Q_i | \rho), \end{aligned}$$

and so

$$M_n^\psi(Q)/n = T_{W_{nj}, \psi}(\mu_n(Q)) - I(\mu_n(Q)). \quad (4.14)$$

As a final preparation for the proof of (4.13), we argue that $\inf_n M_n^\psi/n > -\infty$. To see this, take $B \subset \mathbb{R}$ to be any compact set of positive ρ -measure, and define $\hat{\rho} \ll \rho$ by $d\hat{\rho}/d\rho = 1_B/\rho(B)$. Let $Q_i = \hat{\rho}$ for $i = 1, \dots, n$, and $Q = Q_1 \times \cdots \times Q_n$. Then

$$\frac{1}{n} \sum_{i=1}^n H(Q_i | \rho) = H(\hat{\rho} | \rho) = -\log \rho(B) < \infty,$$

and also

$$T_{W_{nj}, \psi}(\mu_n(Q)) \geq -\|W_{nj}\|_{L_1[0,1]^2} \sup_{x,y \in B} |\psi(x, y)|.$$

Since ψ is continuous, it is bounded on the compact set B . Since W_{nj} converges in strong cut metric to W , we have $\|W_{nj}\|_{L_1[0,1]^2} \rightarrow \|W\|_{L_1[0,1]^2}$, and thus the right-hand side is bounded. This proves that $\inf_n M_n^\psi/n > -\infty$. We now prove the upper and lower bounds in (4.13) separately.

Proof of the upper bound in (4.13): Let $Q^n = Q_1^n \times \cdots \times Q_n^n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ be any near-optimizer of $M_n^\psi(\cdot)$, meaning

$$M_n^\psi(Q^n) \geq M_n^\psi - o(n). \quad (4.15)$$

Note that $M_n^\psi(Q^n)/n$ is bounded from below by some constant C , as shown just above. Since J has nonnegative entries and $\psi \leq 0$, we have $T_{W_{nj}, \psi} \leq 0$ which implies

$$C \leq M_n^\psi(Q^n)/n = T_{W_{nj}, \psi}(\mu_n(Q^n)) - I(\mu_n(Q^n)) \leq -I(\mu_n(Q^n)).$$

This implies $\sup_n I(\mu_n(Q^n)) < \infty$. Since the sub-level sets of I are weakly compact, the sequence $(\mu_n(Q^n))$ has a limit point. Let μ_∞ be any limit point. Lower semicontinuity of $I(\cdot)$ gives $I(\mu_\infty) < \infty$. For each $m \in \mathbb{N}$, define $\psi_m(x, y) := \psi(x, y)1_{\{|x|, |y| \leq m\}}$. Note that $\psi \leq \psi_m \leq 0$, and thus $T_{W_{nj}, \psi} \leq T_{W_{nj}, \psi_m}$. By part (1) of Lemma 4.2,

$$\sup_{\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})} |T_{W_{nj}, \psi_m}(\mu) - T_{W, \psi_m}(\mu)| \rightarrow 0,$$

for all $m \in \mathbb{N}$. Therefore, for all m ,

$$\limsup_{n \rightarrow \infty} T_{W_{nj}, \psi}(\mu_n(Q^n)) \leq \limsup_{n \rightarrow \infty} T_{W_{nj}, \psi_m}(\mu_n(Q^n)) \leq \limsup_{n \rightarrow \infty} T_{W, \psi_m}(\mu_n(Q^n)) = T_{W, \psi_m}(\mu_\infty),$$

where the last step uses part (2) of Lemma 4.2. The left-hand side above does not depend on m , and thus

$$\limsup_{n \rightarrow \infty} T_{W_{nj}, \psi}(\mu_n(Q^n)) \leq \inf_{m \in \mathbb{N}} T_{W, \psi_m}(\mu_\infty) = T_{W, \psi}(\mu_\infty),$$

where the last equality follows from the monotone convergence theorem and the fact that $\psi_m \downarrow \psi$ pointwise. Using the lower semicontinuity of I , we deduce

$$T_{W, \psi}(\mu_\infty) - I(\mu_\infty) \geq \limsup_{n \rightarrow \infty} (T_{W_{nj}, \psi}(\mu_n(Q^n)) - I(\mu_n(Q^n))) = \limsup_{n \rightarrow \infty} M_n^\psi(Q^n)/n.$$

Bound the left-hand side by a supremum to prove the upper bound in (4.13). Moreover, once we prove (4.13), then this argument shows the following: for any near-optimizing sequence $Q^n = \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ in the sense of (4.15), the sequence $\{\mu_n(Q^n)\}$ is tight, and for any limit point μ_∞ of $\{\mu_n(Q^n)\}$ it holds that μ_∞ is an optimizer for the right-hand side of (4.13).

Proof of the lower bound in (4.13): To prove the lower bound in (4.13), we first claim that

$$\sup_{\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})} (T_{W, \psi}(\mu) - I(\mu)) = \sup_{\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R}), \text{ compact support}} (T_{W, \psi}(\mu) - I(\mu)). \quad (4.16)$$

The inequality ($\geq$) is obvious. To prove the reverse, let $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ such that $I(\mu) < \infty$, and define $\mu^m \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ with compact support by setting $d\mu^m/d\mu = 1_{[0, 1] \times [-m, m]}/\mu([0, 1] \times [-m, m])$, which is well defined for large enough m . Then

$$\begin{aligned} I(\mu_m) &= H(\mu^m | \bar{\mu}) = \int_{[0, 1] \times \mathbb{R}} \log \frac{d\mu^m}{d\bar{\mu}} d\mu^m \\ &= \int_{[0, 1] \times \mathbb{R}} \log \frac{d\mu^m}{d\mu} d\mu^m + \int_{[0, 1] \times \mathbb{R}} \log \frac{d\mu}{d\bar{\mu}} d\mu^m. \end{aligned}$$

The second term converges to $I(\mu)$ by dominated convergence. The first term equals $-\log \mu([0, 1] \times [-m, m])$ and vanishes as $m \rightarrow \infty$. Finally, since $W \geq 0$ and $\psi \leq 0$, it is straightforward to check by monotone convergence that $T_{W, \psi}(\mu_m) \rightarrow T_{W, \psi}(\mu)$, and thus $T_{W, \psi}(\mu_m) - I(\mu_m) \rightarrow T_{W, \psi}(\mu) - I(\mu)$ as $m \rightarrow \infty$. This proves (4.16).

Now, to prove the lower bound in (4.13), we let $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ with compact support and with $I(\mu) < \infty$, and note that necessarily $\mu \ll \bar{\mu}$. By defining $h(\cdot, \cdot) := \frac{d\mu}{d\bar{\mu}}$, we have $\int_{\mathbb{R}} h(u, \cdot) d\rho = 1$ for a.e. $u \in [0, 1]$ since both μ and $\bar{\mu}$ have uniform first marginal. For each $i \in [n]$, define $h_i^n : \mathbb{R} \rightarrow [0, \infty)$ by

$$h_i^n(x) := n \int_{(i-1)/n}^{i/n} h(u, x) du.$$

By Fubini's theorem, $\int_{\mathbb{R}} h_i^n d\rho = 1$ for all i . We may thus define $Q^n = Q_1^n \times \cdots \times Q_n^n \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ by setting $\frac{dQ_i^n}{d\rho} = h_i^n$, and define $\mu_n(Q^n)$ as before; note for later use the key identity $\frac{d\mu_n(Q^n)}{d\bar{\mu}}(u, x) = h_{[nu]}^n(x)$. If $\mathcal{K}$

denotes a compact interval such that $[0, 1] \times \mathcal{K}$ contains the support of μ , then $[0, 1] \times \mathcal{K}$ also contains the support of $\mu_n(Q^n)$, and we may replace ψ by $\psi 1_{\mathcal{K}^2}$ in the following argument. Recalling the formula (4.14) for $M_n^\psi(Q)$, we may use part (1) of Lemma 4.2 to get

$$T_{W,\psi}(\mu_n(Q^n)) - \frac{1}{n} M_n^\psi(Q^n) - I(\mu_n(Q^n)) \rightarrow 0. \quad (4.17)$$

To complete the proof of the lower bound, we will show that

$$\lim_{n \rightarrow \infty} T_{W,\psi}(\mu_n(Q^n)) = T_{W,\psi}(\mu), \quad \text{and} \quad I(\mu_n(Q^n)) \leq I(\mu), \quad \forall n. \quad (4.18)$$

Once (4.18) is established, it will follow from the lower semicontinuity of I that $I(\mu_n(Q^n)) \rightarrow I(\mu)$, and we use (4.17) to deduce

$$\liminf_{n \rightarrow \infty} M_n^\psi/n \geq \lim_{n \rightarrow \infty} M_n^\psi(Q^n)/n = T_{W,\psi}(\mu) - I(\mu).$$

This holds for every $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ of compact support satisfying $I(\mu) < \infty$. Hence, taking the supremum and recalling (4.16) yields the desired lower bound in (4.13).

It remains to prove (4.18). Note that

$$\begin{aligned} \int_{[0,1] \times \mathbb{R}} \left| \frac{d\mu}{d\bar{\mu}} - \frac{d\mu_n(Q^n)}{d\bar{\mu}} \right| d\bar{\mu} &= \int_0^1 \int_{\mathbb{R}} |h(u, x) - h_{[nu]}^n(x)| \rho(dx) du \\ &= \mathbb{E}_{\bar{\mu}} |h(U, X) - \mathbb{E}_{\bar{\mu}}[h(U, X)|\mathcal{F}_n]|, \end{aligned}$$

where $\mathcal{F}_n$ is the σ -field generated by $([nU], X)$. The right-hand side converges to 0 by Levy's upwards convergence theorem, since $\mathbb{E}_{\bar{\mu}} |h(U, X)| = 1 < \infty$. Thus the probability measure $\mu_n(Q^n)$ converges in total variation to μ , and the first claim in (4.18) follows from part (2) of Lemma 4.2. To prove the second claim in (4.18), use convexity of $\varphi(x) := x \log x$ for $x \geq 0$, along with Jensen's inequality, to get

$$I(\mu_n(Q^n)) = \mathbb{E}_{\bar{\mu}} \varphi(\mathbb{E}_{\bar{\mu}}[h(U, X)|\mathcal{F}_n]) \leq \mathbb{E}_{\bar{\mu}} \varphi(h(U, X)) = I(\mu).$$

This proves (4.18), completing the proof of the lower bound, and thus Theorem 2.8(1).

Proof of Theorem 2.8(2).

We first discuss the optimization problem. The functional to be optimized can be written as

$$\Phi(\mu) := \frac{1}{2} \int_{[0,1] \times \mathbb{R}} \int_{[0,1] \times \mathbb{R}} W(u, v) K(x - y) \mu(du, dx) \mu(dv, dy) - \int_0^1 H(\mu_u | \rho) du$$

We will show the existence of an optimizer via the weak upper semicontinuity: Since $W \geq 0$ and $K \leq 0$, monotone convergence yields

$$2T_{W,\tilde{K}}(\mu) = \mathbb{E}_{\mu^{\otimes 2}} [W(U_1, U_2)K(X_1 - X_2)] = \inf_{m>0} \mathbb{E}_{\mu^{\otimes 2}} [(W(U_1, U_2)K(X_1 - X_2)) \vee (-m)].$$

For each m , the expectation appearing on the right-hand side is continuous as a function of $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$, by Lemma 4.3. Hence, the left-hand side is upper semicontinuous. Since relative entropy is lower semicontinuous with compact sub-level sets, the existence of an optimizer follows.

We prove uniqueness of the optimizer via displacement convexity. Letting $\tilde{K}(x, y) = (1/2)K(x - y)$, we may rewrite $\Phi(\mu) = \Phi_1(\mu) + T_{W,\tilde{K}}(\mu) - \Phi_2(\mu)$, where we define

$$\Phi_1(\mu) := \int_0^1 \int_{\mathbb{R}} V(x) \mu_u(dx) du, \quad \Phi_2(\mu) := \int_0^1 H(\mu_u) du,$$

where we used the simple identity $H(v|\rho) = H(v) - \int_{\mathbb{R}^n} V dv$. Let $\mu^0, \mu^1 \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$ be two optimizers, written in disintegrated form as $du\mu_u^i(dx)$ for $i = 0, 1$. Let $F_u^i(x) = \mu_u^i(-\infty, x]$ denote the CDF, with generalized inverse $\bar{F}_u^i(y) := \inf\{x \in \mathbb{R} : y \leq F_u^i(x)\}$. Then, for each $u \in [0, 1]$, $T_u(x) := \bar{F}_u^1(F_u^0(x))$ denotes the unique nondecreasing function with $(T_u)_\# \mu_u^0 = \mu_u^1$. Since $F_u^i(x)$ is right-continuous in x and measurable in u , it is jointly measurable in (u, x) , and the same is easily seen to be true for $\bar{F}_u^i(x)$ and thus $T_u(x)$. Consider the map $\bar{T} : [0, 1] \times \mathbb{R} \rightarrow [0, 1] \times \mathbb{R}$ given by $\bar{T}(u, x) = (u, T_u(x))$. Define the interpolation $\mu^t := ((1-t)\text{Id} + t\bar{T})_\# \mu^0$ for each $t \in [0, 1]$. Then we have

$$\begin{aligned} T_{W, \bar{K}}(\mu^t) &= \frac{1}{2} \int_{[0,1] \times \mathbb{R}} \int_{[0,1] \times \mathbb{R}} W(u, v) K(x - y) \mu^t(du, dx) \mu^t(dv, dy) \\ &= \frac{1}{2} \int_{[0,1] \times \mathbb{R}} \int_{[0,1] \times \mathbb{R}} W(u, v) K((1-t)(x-y) + t(T_u(x) - T_u(y))) \mu^0(du, dx) \mu^0(dv, dy). \end{aligned}$$

Since K is concave and $W \geq 0$, $t \mapsto T_{W, \bar{K}}(\mu^t)$ is concave. Note also that

$$\Phi_2(\mu^t) = \int_0^1 H(((1-t)\text{Id} + tT_u)_\# \mu_u^0) du$$

is a convex function of t , by the displacement convexity of entropy [57, Theorem 5.15(i)]. By the κ -concavity of V , the function $t \mapsto \Phi_1(\mu^t)$ is strictly concave, and we find that $t \mapsto \Phi(\mu^t)$ is strictly concave. Since μ^0 and μ^1 are both optimizers, we have $\Phi(\mu^0) = \Phi(\mu^1)$. Hence, we must have $\mu^0 = \mu^1$, as otherwise the strict concavity would be contradicted.

With existence and uniqueness of the optimizer settled, we lastly prove the claim (2.4) in part (2.8) of Theorem 2.8. Note that Theorem 1.1 implies uniqueness of the optimizer Q^n in $\sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} M_n(Q)$ for each n . Since Q^n is optimal and thus a fortiori near-optimal, we may use the following fact proven in the course of proving the upper bound in Theorem 2.8(1): The sequence $\{\mu_n(Q^n)\}$ is tight (since Q^n is), and any limit point is an optimizer for the right-hand side of (4.13). We have just shown the latter optimizer to be unique, and let us denote it $\mu^* \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R})$. Thus, $\mu_n(Q^n) \rightarrow \mu^*$ weakly. From part (1) and Corollary 1.2, for any bounded 1-Lipschitz function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ we have

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\mathcal{P}} \left[\left(\frac{1}{n} \sum_{i=1}^n \varphi(X_i) - \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{Q^n}[\varphi(X_i)] \right)^2 \right] = 0.$$

Note that

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}_{Q^n}[\varphi(X_i)] = \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} \varphi(x) Q_i^n(dx) = \int_{[0,1] \times \mathbb{R}} \varphi(x) \mu_n(Q^n)(du, dx).$$

Using the weak convergence $\mu_n(Q^n) \rightarrow \mu^*$, the right-hand side converges to

$$\int_{[0,1] \times \mathbb{R}} \varphi(x) \mu^*(du, dx) = \int_{\mathbb{R}} \varphi dR^*, \quad \text{where } R^* := \int_0^1 \mu_u^* du.$$

We deduce that $\frac{1}{n} \sum_{i=1}^n \varphi(X_i) \rightarrow \int_{\mathbb{R}} \varphi dR^*$ in probability for each bounded Lipschitz φ . This is enough to deduce the convergence in distribution $\frac{1}{n} \sum_{i=1}^n \delta_{X_i} \rightarrow R^*$.

4.3 Proof of Lemma 2.11

We first prove (1). When f is even, we claim that (the density of) Q^* is also even, which completes the proof because it implies $\mathbb{E}_{Q^*}[X_i] = 0$ for all i . To show that Q^* is even, let $R_i(x) := Q_i^*(-x)$ for each $x \in \mathbb{R}$ and $i = 1, \dots, n$. Let $R = R_1 \times \dots \times R_n$. Then $\int_{\mathbb{R}^n} f dR = \int_{\mathbb{R}^n} f dQ$ by evenness of f , and clearly $H(Q) = H(R)$. Hence, R is also an optimizer of (1.4), and we deduce $R = Q^*$ by uniqueness of the optimizer.

We prove (2) by showing in this case that $Q_i^* = Q_j^*$ for all i, j . Suppose f is invariant with respect to a transitive set S of permutations of $[n]$. Fix $i, j \in \{1, \dots, n\}$. Choose $\pi \in S$ such that $\pi(i) = j$, which is possible by the assumed transitivity of S . Let $R_k = Q_{\pi(k)}^*$ for each $k = 1, \dots, n$, and let $R = R_1 \times \dots \times R_n$.

The invariance of f under S ensures that $\int_{\mathbb{R}^n} f dR = \int_{\mathbb{R}^n} f dQ^*$. Clearly, $H(R) = H(Q^*)$. Hence, R is also an optimizer of (1.4), and we deduce that $R = Q^*$ by uniqueness. Since $\pi(i) = j$, this implies $Q_i^* = R_i = Q_j^*$.

5 Stochastic Control Proofs

As explained in Remark 2.15, the optimal admissible pair (α, X) for (2.7) is given by

$$\alpha_g(t, x) = \nabla_x \log \mathbb{E}[e^{ng(x+B_T-B_t)}], \quad (5.1)$$

with $X = (X_t)_{t \in [0, T]}$ being the Brownian bridge with terminal law $P(dx) = Z^{-1} e^{ng(x)} \gamma_T(dx)$. Letting $\mathbb{P}$ denote the Wiener measure on $C([0, T]; \mathbb{R}^n)$, the law $\mathbb{Q}^P$ of this process X can be characterized as the unique minimizer of $\mathbb{Q} \mapsto H(\mathbb{Q} | \mathbb{P})$ among $\mathbb{Q}$ with time- T marginal equal to P ; see [9, Proposition 6] or [47, Lemma 10]. This minimizer satisfies

$$H(\mathbb{Q}^P | \mathbb{P}) = H(P | \gamma_T) = \frac{1}{2} \mathbb{E} \left[\int_0^T |\alpha_g(t, X_t)|^2 dt \right]. \quad (5.2)$$

Note that $H(P | \gamma_T) < \infty$, and so the pair (α_g, X) is admissible in the sense of Section 2.3.

Proof of Corollary 2.14. Once the formulas (2.10) and (2.12) are established, the final claim follows immediately from Corollary 1.4, applied with $V_i(x) = -x^2/(2T)$ for $i = 1, \dots, n$ and $\kappa = 1/T$.

To prove (2.10) and (2.12), we begin with the inequality $(\leq)$. Let (α, X) denote any admissible pair, and let $\mathbb{Q}$ denote the law of $X = (X_t)_{t \in [0, T]}$. A well known argument using Girsanov's theorem [47, Proposition 1] yields

$$H(\mathbb{Q} | \mathbb{P}) \leq \frac{1}{2} \mathbb{E} \int_0^T |\alpha(t, X_t)|^2 dt = \frac{1}{2} \sum_{i=1}^n \mathbb{E} \int_0^T |\alpha_i(t, X_t)|^2 dt.$$

With $\mathbb{Q}_T$ denoting the law of X_T , note that marginalizing (at time T) does not increase entropy: $H(\mathbb{Q} | \mathbb{P}) \geq H(\mathbb{Q}_T | \gamma_T)$. Thus,

$$\begin{aligned} \mathbb{E} \left[g(X_T) - \frac{1}{2n} \sum_{i=1}^n \int_0^T |\alpha_i(t, X_t)|^2 dt \right] &\leq \int_{\mathbb{R}^n} g d\mathbb{Q}_T - \frac{1}{n} H(\mathbb{Q}_T | \gamma_T) \\ &\leq \sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - \frac{1}{n} H(Q | \gamma_T) \right). \end{aligned}$$

Taking a supremum over all admissible pairs (α, X) proves the inequality $(\leq)$ in (2.10). Now, if (α, X) is an *distributed* admissible pair, then the same chain of inequalities holds, but also $\mathbb{Q}_T$ is a product measure. We can thus deduce (2.12) in the same manner.

The inequality $(\geq)$ in (2.10) and (2.12) follows quickly from the entropy identity (5.2). Starting with (2.10), let $X = (X_t)_{t \in [0, T]}$ be the Brownian bridge with terminal law $P(dx) = Z^{-1} e^{ng(x)} \gamma_T(dx)$. Let α_g be given as in (5.1). By the Gibbs variational principle [33, Proposition 1.4.2], the supremum in (2.10) is attained by $Q = P$. Using $X_T \sim P$ and (5.2), we obtain

$$\begin{aligned} \sup_{Q \in \mathcal{P}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g dQ - \frac{1}{n} H(Q | \gamma_T) \right) &= \int_{\mathbb{R}^n} g dP - \frac{1}{n} H(P | \gamma_T) \\ &= \mathbb{E} \left[g(X_T) - \frac{1}{2n} \int_0^T |\alpha_g(t, X_t)|^2 dt \right] \leq V_{\text{orig}}. \end{aligned}$$

This proves $(\geq)$ in (2.10), and also proves that (α_g, X) is optimal. Similarly, to prove the inequality $(\geq)$ in (2.12), let $Q^* \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)$ be the unique optimizer in (2.12), which we know by Corollary 1.4 to take the form stated in Corollary 2.14. Let $X = (X_t)_{t \in [0, T]}$ be the Brownian bridge with terminal law Q^* . Define α_g

as in (5.1), with $\hat{g}(x) = \sum_{i=1}^n \mathbb{E}_{Q^*}[g(X)|X_i = x_i]$ in place of g . Using $X_T \sim Q^*$ and (5.2), we obtain

$$\begin{aligned} \sup_{Q \in \mathcal{P}_{\text{pr}}(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} g \, dQ - \frac{1}{n} H(Q | \gamma_T) \right) &= \int_{\mathbb{R}^n} g \, dQ^* - \frac{1}{n} H(Q^* | \gamma_T) \\ &= \mathbb{E} \left[g(X_T) - \frac{1}{2n} \int_0^T |\alpha_{\hat{g}}(t, X_t)|^2 \, dt \right] \leq V_{\text{distr}}. \end{aligned}$$

Indeed, note that $(\alpha_{\hat{g}}, X)$ is an admissible distributed pair because Q^* is a product measure. This proves $(\geq)$ in (2.12), and also proves that $(\alpha_{\hat{g}}, X)$ is optimal. ■

Funding

D.L. and L.C.Y. are partially supported by the AFOSR Grant FA9550-19-1-0291 and the NSF CAREER award DMS-2045328. S.M. is partially supported by NSF grants DMS-1712037 and DMS-2113414.

Acknowledgments

We thank Ronen Eldan for helpful discussions and comments.

1. Agueh, M. and G. Carlier. “Barycenters in the wasserstein space.” *SIAM J. Math. Anal.* **43**, no. 2 (2011): 904–24. <https://doi.org/10.1137/100805741>.
2. Alquier, P., J. Ridgway, and N. Chopin. “On the properties of variational approximations of Gibbs posteriors.” *J. Mach. Learn. Res.* **17**, no. 1 (2016): 8374–414.
3. Augeri, F. “Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős-Rényi graphs.” *Ann. Probab.* **48**, no. 5 (2020): 2404–48. <https://doi.org/10.1214/20-AOP1427>.
4. Augeri, F. “A transportation approach to the mean-field approximation.” *Probab. Theory Related Fields* **180**, no. 1 (2021): 1–32. <https://doi.org/10.1007/s00440-021-01056-2>.
5. Austin, T. “The structure of low-complexity Gibbs measures on product spaces.” *Ann. Probab.* **47**, no. 6 (2019): 4002–23. <https://doi.org/10.1214/19-AOP1352>.
6. Bakry, D. and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, pp. 177–206. Springer, 1985, <https://doi.org/10.1007/BFb0075847>.
7. Bakry, D., I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, Vol. 348. Springer Science & Business Media, 2013, <https://doi.org/10.1007/978-3-319-00227-9>.
8. Basak, A. and S. Mukherjee. “Universality of the mean-field for the Potts model.” *Probab. Theory Related Fields* **168**, no. 3–4 (2017): 557–600. <https://doi.org/10.1007/s00440-016-0718-0>.
9. Baudoin, F. “Conditioned stochastic differential equations: theory, examples and application to finance.” *Stochastic Process. Appl.* **100**, no. 1–2 (2002): 109–45. [https://doi.org/10.1016/S0304-4149\(02\)00109-6](https://doi.org/10.1016/S0304-4149(02)00109-6).
10. Biskup, M. and L. Chayes. “Rigorous analysis of discontinuous phase transitions via mean-field bounds.” *Comm. Math. Phys.* **238**, no. 1 (2003): 53–93. <https://doi.org/10.1007/s00220-003-0828-2>.
11. Blei, D. M., A. Kucukelbir, and J. D. McAuliffe. “Variational inference: a review for statisticians.” *J. Amer. Statist. Assoc.* **112**, no. 518 (2017): 859–77. <https://doi.org/10.1080/01621459.2017.1285773>.
12. Bobkov, S. and M. Madiman. “The entropy per coordinate of a random vector is highly constrained under convexity conditions.” *IEEE Trans. Inform. Theory* **57**, no. 8 (2011): 4940–54. <https://doi.org/10.1109/TIT.2011.2158475>.
13. Bobkov, S. G. and F. Götze. “Exponential integrability and transportation cost related to logarithmic Sobolev inequalities.” *J. Funct. Anal.* **163**, no. 1 (1999): 1–28. <https://doi.org/10.1006/jfan.1998.3326>.
14. Borell, C. “Diffusion equations and geometric inequalities.” *Potential Anal.* **12**, no. 1 (2000): 49–71. <https://doi.org/10.1023/A:1008641618547>.
15. Borgs, C., J. T. Chayes, H. Cohn, and Y. Zhao. “An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence.” *Ann. Probab.* **46**, no. 1 (2018): 337–96.
16. Borgs, C., J. T. Chayes, H. Cohn, and Y. Zhao. “An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions.” *Trans. Amer. Math. Soc.* **372**, no. 5 (2019): 3019–62.

17. Borgs, C., J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztegombi. "Convergent sequences of dense graphs I: subgraph frequencies, metric properties and testing." *Adv. Math.* **219**, no. 6 (2008): 1801–51. <https://doi.org/10.1016/j.aim.2008.07.008>.
18. Boué, M. and P. Dupuis. "A variational representation for certain functionals of Brownian motion." *Ann. Probab.* **26**, no. 4 (1998): 1641–59. <https://doi.org/10.1214/aop/1022855876>.
19. Brascamp, H.J. and E.H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. In *Inequalities*, pp. 441–464. Springer, 2002, https://doi.org/10.1007/978-3-642-55925-9_36.
20. Carmona, R. and F. Delarue. "Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics." *Ann. Probab.* **43**, no. 5 (2015): 2647–700. <https://doi.org/10.1214/14-AOP946>.
21. Carmona, R. and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018, <https://doi.org/10.1007/978-3-319-56436-4>.
22. Carrillo, J. A., R. J. McCann, and C. Villani. "Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates." *Rev. Mat. Iberoam.* **19**, no. 3 (2003): 971–1018.
23. Chafaï, D., N. Gozlan, and P.-A. Zitt. "First-order global asymptotics for confined particles with singular pair repulsion." *Ann. Appl. Probab.* **24**, no. 6 (2014): 2371–413. <https://doi.org/10.1214/13-AAP980>.
24. Chatterjee, S. and A. Dembo. "Nonlinear large deviations." *Adv. Math.* **299** (2016): 396–450. <https://doi.org/10.1016/j.aim.2016.05.017>.
25. Cook, N. and A. Dembo. "Large deviations of subgraph counts for sparse Erdős-Rényi graphs." *Adv. Math.* **373** (2020): 107289. <https://doi.org/10.1016/j.aim.2020.107289>.
26. Cook, N. A., A. Dembo, and H.-T. Pham. "Regularity method and large deviation principles for the Erdős-Rényi hypergraph." arXiv preprint, arXiv:2102.09100, 2021.
27. Csiszár, I. "I-divergence geometry of probability distributions and minimization problems." *Ann. Probab.* **3** (1975): 146–58. <https://doi.org/10.1214/aop/1176996454>.
28. Deb, N. and S. Mukherjee. "Fluctuations in mean-field Ising models." *Ann. Appl. Probab.* **33**, no. 3 (2023): 1961–2003. <https://doi.org/10.1214/22-AAP1857>.
29. Delattre, S., G. Giacomini, and E. Luçon. "A note on dynamical models on random graphs and Fokker-Planck equations." *J. Statist. Phys.* **165**, no. 4 (2016): 785–98. <https://doi.org/10.1007/s10955-016-1652-3>.
30. Dembo, A. and A. Montanari. "Gibbs measures and phase transitions on sparse random graphs." *Braz. J. Probab. Stat.* **24**, no. 2 (2010): 137–211. <https://doi.org/10.1214/09-BJPS027>.
31. Dembo, A. and A. Montanari. "Ising models on locally tree-like graphs." *Ann. Appl. Probab.* **20**, no. 2 (2010): 565–92. <https://doi.org/10.1214/09-AAP627>.
32. Dembo, A., A. Montanari, A. Sly, and N. Sun. "The replica symmetric solution for Potts models on d-regular graphs." *Comm. Math. Phys.* **327**, no. 2 (2014): 551–75. <https://doi.org/10.1007/s00220-014-1956-6>.
33. Dupuis, P. and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*, Vol. 902. John Wiley & Sons, 2011.
34. Dupuis, P., V. Laschos, and K. Ramanan. "Large deviations for configurations generated by Gibbs distributions with energy functionals consisting of singular interaction and weakly confining potentials." *Electron. J. Probab.* **25** (2020): 1–41. <https://doi.org/10.1214/20-EJP449>.
35. Eldan, R. "Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations." *Geom. Funct. Anal.* **28**, no. 6 (2018): 1548–96. <https://doi.org/10.1007/s00039-018-0461-z>.
36. Eldan, R. "Taming correlations through entropy-efficient measure decompositions with applications to mean-field approximation." *Probab. Theory Related Fields* **176**, no. 3 (2020): 737–55. <https://doi.org/10.1007/s00440-019-00924-2>.
37. Eldan, R. and R. Gross. "Decomposition of mean-field Gibbs distributions into product measures." *Electron. J. Probab.* **23** (2018). <https://doi.org/10.1214/18-EJP159>.
38. Ellis, R. S. and C. M. Newman. "The statistics of Curie-Weiss models." *J. Statist. Phys.* **19**, no. 2 (1978): 149–61. <https://doi.org/10.1007/BF01012508>.
39. Fleming, W. H. and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*, Vol. 25. Springer Science & Business Media, 2006.
40. Föllmer, H. An entropy approach to the time reversal of diffusion processes. In *Stochastic Differential Systems: Filtering and Control*, pp. 156–163. Springer, 1985, <https://doi.org/10.1007/BFb0005070>.
41. Föllmer, H. Time reversal on Wiener space. In *Stochastic Processes-Mathematics and Physics*, pp. 119–129. Springer, 1986, <https://doi.org/10.1007/BFb0080212>.

42. Huang, M., R. Malhamé, and P. Caines. "Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the Nash certainty equivalence principle." *Commun. Inf. Syst.* **6**, no. 3 (2006): 221–52.
43. Jackson, J. and D. Lacker. "Approximately optimal distributed stochastic controls beyond the mean field setting." arXiv preprint, arXiv:2301.02901. (2023).
44. Jain, V., F. Koehler, and A. Risteski. "Mean-field approximation, convex hierarchies, and the optimality of correlation rounding: a unified perspective." *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, 1226–36, 2019.
45. Lacker, D., K. Ramanan, and R. Wu. "Local weak convergence for sparse networks of interacting processes." *Ann. Appl. Probab.* **33**, no. 2 (2023): 843–88. <https://doi.org/10.1214/22-AAP1830>.
46. Lasry, J.-M., and P.-L. Lions "Mean field games." *Japan. J. Math. (N.S.)* **2**, no. 1 (2007): 229–60. <https://doi.org/10.1007/s11537-007-0657-8>.
47. Lehec, J. "Representation formula for the entropy and functional inequalities." *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* **49**, no. 3 (2013): 885–99. <https://doi.org/10.1214/11-AIHP464>.
48. Lovász, L. *Large Networks and Graph Limits*, vol. 60. American Mathematical Soc., 2012, <https://doi.org/10.1090/coll/060>.
49. Lubetzky, E. and Y. Zhao. "On the variational problem for upper tails in sparse random graphs." *Random Struct. Algorithms* **50**, no. 3 (2017): 420–36. <https://doi.org/10.1002/rsa.20658>.
50. McCann, R. J. "A convexity principle for interacting gases." *Adv. Math.* **128**, no. 1 (1997): 153–79. <https://doi.org/10.1006/aima.1997.1634>.
51. Mukherjee, S. and S. Sen. "Variational inference in high-dimensional linear regression." *J. Mach. Learn. Res.* **23**, no. 1 (2022): 13703–58.
52. Oliveira, R. and G. Reis. "Interacting diffusions on random graphs with diverging average degrees: hydrodynamics and large deviations." *J. Statist. Phys.* **176**, no. 5 (2019): 1057–87. <https://doi.org/10.1007/s10955-019-02332-1>.
53. Otto, F. and C. Villani. "Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality." *J. Funct. Anal.* **173**, no. 2 (2000): 361–400. <https://doi.org/10.1006/jfan.1999.3557>.
54. Ray, K. and B. Szabó. "Variational bayes for high-dimensional linear regression with sparse priors." *J. Amer. Statist. Assoc.* (2021): 1–12.
55. Revuz, D. and M. Yor. *Continuous Martingales and Brownian Motion*, Vol. 293. Springer Science & Business Media, 2013.
56. Saumard, A. and J. A. Wellner. "Log-concavity and strong log-concavity: a review." *Stat. Surv.* **8** (2014): 45–114.
57. Villani, C. *Topics in Optimal Transportation*, 58. American Mathematical Society, 2003, <https://doi.org/10.1090/gsm/058>.
58. Wainwright, M. J. and M. I. Jordan. "Graphical models, exponential families, and variational inference." *Found. Trends Mach. Learn.* **1**, no. 1–2 (2008): 1–305. <https://doi.org/10.1561/22000000001>.
59. Yan, J. "Nonlinear large deviations: beyond the hypercube." *Ann. Appl. Probab.* **30**, no. 2 (2020): 812–46.
60. Zhang, A. Y. and H. H. Zhou. "Theoretical and computational guarantees of mean field variational inference for community detection." *Ann. Stat.* **48**, no. 5 (2020): 2575–98. <https://doi.org/10.1214/19-AOS1898>.