Mean-field limit for a class of stochastic ergodic control problems

Sergio Albeverio[1,*]**, Francesco C. De Vecchi**[1,**]**, Andrea Romano**[2,†] **and Stefania Ugolini**[2,‡]

[1]*Institute for Applied Mathematics and Hausdorff Center for Mathematics, Endenicher Allee 60, 53115 Bonn, Germany. e-mail:*
[*]`albeverio@iam.uni-bonn.de`*;* [**]`francesco.devecchi@uni-bonn.de`

[2]*Dipartimento di Matematica, Via Saldini 50, 20113 Milano, Italy. e-mail:*
[†]`andrea.romano4@studenti.unimi.it`*;* [‡]`stefania.ugolini@unimi.it`

Abstract: We study a family of McKean-Vlasov (mean-field) type ergodic optimal control problems with linear control, and quadratic dependence on control of the cost function. For this class of problems we establish existence and uniqueness of an optimal control. We propose an N-particles Markovian optimal control problem approximating the McKean-Vlasov one and we prove the convergence in relative entropy, total variation and Wasserstein distance of the law of the former to the law of the latter when N goes to infinity. Some McKean-Vlasov optimal control problems with singular cost function and the relation of these problems with the mathematical theory of Bose-Einstein condensation is also established.

Keywords and phrases: mean field control, ergodic optimal control, McKean–Vlasov limit, de Finetti theorem, strong Kac's chaos, convergence of probability measures on path space, singular cost functional.

1. Introduction

In this paper we want to provide a complete study of a family of mean-field ergodic stochastic optimal control problems, known as optimally controlled McKean-Vlasov dynamics. More precisely we consider the controlled stochastic differential equation (SDE)

$$dX_t = \alpha(X_t)dt + \sqrt{2}dW_t \tag{1.1}$$

where α is a C^1 control function from $\mathbb{R}^n$ to $\mathbb{R}^n$ and W_t, $t \geq 0$, is an n dimensional standard Brownian motion, with the following cost functional

$$J(\alpha, x_0) = \limsup_{T \to +\infty} \frac{1}{T} \left(\int_0^T \mathbb{E}_{x_0} \left[\frac{|\alpha(X_t)|^2}{2} + \mathcal{V}(X_t, \mathrm{Law}(X_t)) \right] dt \right). \tag{1.2}$$

Here $\mathcal{V} : \mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}$ (where $\mathcal{P}(\mathbb{R}^n)$ is the space of probability measures on $\mathbb{R}^n$ endowed with the metric given by the weak convergence) is a regular function satisfying some technical hypotheses (see Section 2 below) and $\mathbb{E}_{x_0}$ is the expectation with respect to the solution X_t to the SDE (1.1) such that $X_0 = x_0 \in \mathbb{R}^n$. We prove existence and uniqueness of the optimal control $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ for the problem given by (1.1) and (1.2). Furthermore we present an N-particle Markovian approximation of the previous problem and give a proof of the convergence of the corresponding value functions and of the invariant finite dimensional probability laws to those given by the one of the McKean-Vlasov dynamics. We remark that the cost functional explicitly depends on the law of X_t and, under natural assumptions, we prove that the optimal control can also be expressed in terms of the same law.

Recently there has been a growing interest in optimally controlled McKean-Vlasov dynamics (see, for example, [11; 12; 13; 22; 23; 62; 61]). The main part ot the current literature focuses on finite or infinite time horizon problems and usually does not discuss the approximation of the controlled McKean-Vlasov problem by Markovian controlled N-particle systems. To the best of our knowledge some of the few exceptions are [24, Chapter 6] (see also [22]), where, for the case of controlled McKean-Vlasov dynamics, the convergence of the value function is considered (which implies the convergence of the optimal trajectory see again [24, Chapter 6]), and [44] and [32; 33]. In particular [44] studies the convergence problem under general conditions, without symmetry assumptions and in the time-dependent setting using a martingale problem approach. In [32] the techniques of [44] are generalized to the case of controlled McKean-Vlasov problems having generally dependent noises, and in [33] a propagation of chaos result is proven for *extended* mean field control problems.
The optimally controlled McKean-Vlasov dynamics is closely related to mean-field games theory (see [23, Chapter 6] for a discussion about the relation between the two approaches). Mean-field games theory, in the case of a finite and infinite time horizon utility function, is much more developed both in the study of the limit problem and in the study of N-particles approximations (see, e.g., the books [16; 23; 24] and references therein as well

[40]). The PDE system related to the ergodic mean-field games is well studied (see, e.g., [15; 17; 18; 27; 46]). In the mean-field games case, the ergodic stochastic problem is considered in [6; 9; 10; 34], see also [19].

Our convergence scheme is quite different from the one usually formulated in the literature on competitive mean-field games, where the value function can be decomposed into the product of the one-particle marginals (see [6; 46]). Indeed our N-particles process is an interacting controlled diffusions system where the chaoticity property is achieved only asymptotically (that is in the infinite particles limit). It is important to note that, as it is proved in [19], in the ergodic competitive mean-field game case, when the control depends on all N-players, the convergence of the N-particle system to the mean-field one in general does not hold. This is one of the main differences with respect to the (cooperative) McKean-Vlasov systems treated in the present paper. To the best of our knowledge, this is the first paper facing in an ergodic framework the convergence problem of a Markovian N interacting diffusions system to a Markovian limit system of McKean-Vlasov type.

The main idea of the paper is to exploit some methods of mathematical physics, in particular from the mathematical theory of Bose-Einstein condensation (see, e.g.[49; 50; 52; 53; 54; 59; 65; 66]), Nelson's stochastic mechanics (see, e.g., [20; 21; 38; 60; 75]), and variational stochastic processes connected with Schrödinger problem in optimal transport and the Hopf-Cole transformation (see, e.g., [8; 25; 28; 47; 48; 56; 77] see also [27; 73]). Our method uses however a compactness argument, hence it does not yield per se convergence rates of optimal trajectories (such convergence results exist in other settings, but under the stronger assumption of convexity, see [22; 23]).

The paper contains three main results. The first one is the proof of existence and uniqueness of the optimal control for the SDE (1.1) with cost functional (1.2) under the technical Hypotheses $\mathcal{V}$ (concerning the functional $\mathcal{V}$ in equation (1.2)) and a convexity request C$\mathcal{V}$ for the cost functional discussed in Section 2. We show that some results ([14]) guarantee that under appropriate conditions the control term is the logarithmic derivative of the probability density of the process and so the cost functional can be expressed in terms of the process probability density. By exploiting calculus of variations we then provide a necessary condition for the optimality of the process probability density (Theorem 17). The method applied here makes no direct use of Hamilton-Jacobi-Bellman equation.

Our second main result consists in the convergence of the value function (or rather the constant which gives the value of the cost functional evaluated at the optimal control) and of the finite dimensional invariant distributions of the Markovian N-particle approximation to the one of the McKean-Vlasov optimal control problem when the number of particles N tends to $+\infty$ (see Theorem 25, Theorem 37 and Remark 38). The convergence of the value function, under the previous Hypotheses $\mathcal{V}$ and C$\mathcal{V}$, is achieved using in an essential way de Finetti theorem for exchangeable particles and some important properties of Fisher information (see Section 5). The convergence of the finite dimensional distributions, under the additional quadratic growth Hypothesis Q$\mathcal{V}$ (see Remark 44 for an analysis of the role of Hypothesis Q$\mathcal{V}$ in the proof of Theorem 37), is obtained by proving that the relative entropy between the finite dimensional distributions, converges to 0 for $N \to +\infty$ (Theorem 37). The latter is achieved by exploiting some results from the theory of interacting particles systems and optimal transport (see [39]).

The third main result is the convergence in total variation of the law of the N-particles approximation to the law of the McKean-Vlasov system. In this way we also establish (see Theorem 46) that the strong Kac's chaos holds for the probability law of the N-interacting controlled diffusions system in the limit of infinitely many particles (in the sense of [45]). Namely we prove that, under Hypotheses $\mathcal{V}$, for any $k > 0$, if $\mathbb{P}_{0,N}^{(k)}$ is the law (on path space) of the first k particles of the optimal N-particles approximation and $\mathbb{P}_0$ is the law (on path space) of the optimal McKean-Vlasov system we have $\mathcal{H}(\mathbb{P}_{0,N}^{(k)}|\mathbb{P}_0^{\otimes k}) \to 0$ as $N \to +\infty$ (where $\mathcal{H}(\cdot|\cdot)$ denotes the relative entropy of the first measure with respect to the second one). This kind of convergence implies the convergence in total variation (Corollary 48) and in Wasserstein metric W_p (for $1 \le p \le 2$) (Corollary 50). Let us remark that the result that is most closely related to our own are in [32; 44], where a more general problem is treated in the finite time horizon case proving a convergence in Wasserstein metric. The ergodic case treated and the type of convergence proved in our paper are however new. Indeed, for example, differently from [44], we start the system at the invariant measure, we have to prove that the N-particles invariant measures converge to the limit one. Furthermore we think that the entropy convergence does not hold in the general setting considered in [44], since when the noise is multiplicative, with diffusion coefficients depending on the control or the law of the solution process, the law of the N-particles approximation and of the limit solution are not longer mutually absolutely continuous.

The plan of the paper is as follows. In Section 2 we define our class of ergodic McKean-Vlasov optimal stochastic control problems, making explicit all our hypotheses and providing a non trivial family of cost functions satisfying them. In Section 3 we prove existence and uniqueness of the optimal control for our problem.

In Section 4 we introduce the Markovian N-particles controlled system used to approximate the McKean-Vlasov dynamics. In Section 5 we prove the convergence of the value function of the N-particles approximation to the one of the McKean-Vlasov problem, and in Section 6 the convergence of the probability law at fixed time is discussed. In Section 7 the process convergence result on the path space in the infinite particles limit is established. A comparison with the mathematical physics literature and a comment on the result for the case of singular potentials are provided in Section 8.

2. The setting and the hypotheses

We consider controlled SDE given by (1.1), with the assumptions stated there. Thus $X = (X_t)_{t\geq 0}$ is an n dimensional process, W is an n dimensional Brownian motion and $\alpha(X_t)$ is the control process, with $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ a C^1 function. We denote by $L_\alpha = \Delta + \alpha \cdot \nabla$ the generator associated with the equation (1.1) and by L_α^* the adjoint of L_α with respect to the Lebesgue measure.

We take a functional
$$\mathcal{V} : \mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n) \to \mathbb{R},$$
where $\mathcal{P}(\mathbb{R}^n)$ is the set of probability measures on $\mathbb{R}^n$. We also define for any $\mu \in \mathcal{P}(\mathbb{R}^n)$ (such that $\mathcal{V}(\cdot, \mu)$ if μ integrable)
$$\tilde{\mathcal{V}}(\mu) := \int_{\mathbb{R}^n} \mathcal{V}(x, \mu)\mu(dx). \tag{2.1}$$

If $\mathcal{K} : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}$ is a function we say that $\mathcal{K}$ is Gâteaux differentiable if for any $\mu, \mu' \in \mathcal{P}(\mathbb{R}^n)$ there exists a bounded continuous function $\partial_\mu \mathcal{K}(\cdot, \mu) : \mathbb{R}^n \to \mathbb{R}$ such that
$$\lim_{\epsilon \to 0^+} \frac{\mathcal{K}(\mu + \epsilon(\mu' - \mu)) - \mathcal{K}(\mu)}{\epsilon} = \int_{\mathbb{R}^n} \partial_\mu \mathcal{K}(y, \mu)(\mu'(dy) - \mu(dy)), \tag{2.2}$$

(the left hand side of (2.2) is well defined, when the limit is defined, since, when $\epsilon \geq 0$, $\mu + \epsilon(\mu' - \mu)$ is a probability measure being the convex combination of two probability measures). Since the function $(\partial_\mu \mathcal{K})(y, \mu)$ is only uniquely determined up to a constant (since $\int_{\mathbb{R}^n}(\mu'(dy) - \mu(dy)) = 0$ being μ and μ' two probability measures), we can choose the normalization condition given by
$$\int_{\mathbb{R}^n} (\partial_\mu \mathcal{K})(y, \mu)\mu(dy) = 0.$$

If a function $\bar{\mathcal{K}} : \mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}$ depends also on $x \in \mathbb{R}^n$ we say that $\bar{\mathcal{K}}$ is Gâteaux differentiable if $\bar{\mathcal{K}}(x, \cdot)$ is Gâteaux differentiable for any $x \in \mathbb{R}^n$. In this case we write
$$\lim_{\epsilon \to 0^+} \frac{\bar{\mathcal{K}}(x, \mu + \epsilon(\mu' - \mu)) - \bar{\mathcal{K}}(x, \mu)}{\epsilon} = \int_{\mathbb{R}^n} \partial_\mu \bar{\mathcal{K}}(x, y, \mu)(\mu'(dy) - \mu(dy)).$$

We formulate the following hypotheses on $\mathcal{V}$ (see (1.2)):

- **Hypotheses $\mathcal{V}$:**
    - *i The map $\mathcal{V}$ is continuous from $\mathbb{R}^n \times \mathcal{P}(\mathbb{R}^n)$ to $\mathbb{R}$ (where $\mathcal{P}(\mathbb{R}^n)$ is equipped with the weak topology of convergence of measures).*
    - *ii There is a positive function V such that*
    $$|\partial^\alpha V(x)| \leq C_\alpha V(x) \qquad V(x) \leq C_1 V(y) \exp(C_2|x - y|), \ x, y \in \mathbb{R}^n \tag{2.3}$$
    where $\alpha \in \mathbb{N}^n$ is a multiindex of length at most $|\alpha| \leq 2$, C_α, C_1 and C_2 are positive constants, and growing to $+\infty$ as $|x| \to +\infty$. Furthermore there are three positive constants c_1, c_2, c_3, with $c_2 > 0$, such that for any $\mu \in \mathcal{P}(\mathbb{R}^n)$:
    $$V(x) - c_1 \leq \mathcal{V}(x, \mu) \leq c_2 V(x) + c_3 \ x \in \mathbb{R}^n. \tag{2.4}$$
    - *iii The map $\mathcal{V}$ is Gâteaux differentiable and $\partial_\mu \mathcal{V}(x, y, \mu)$ is uniformly bounded from below and we have*
    $$\partial_\mu \mathcal{V}(x, y, \mu) \leq D_1 + D_2 V(x)V(y), \ x, y \in \mathbb{R}^n \tag{2.5}$$

for some $D_1, D_2 \geq 0$. Furthermore whenever

$$\partial_\mu \tilde{\mathcal{V}}(y, \mu) = \mathcal{V}(y, \mu) + \int_{\mathbb{R}^n} \partial_\mu \mathcal{V}(x, y, \mu) \mu(dx)$$

is well defined (namely when $\int_{\mathbb{R}^n} V(x)\mu(dx) < +\infty$), we require that $\partial_\mu \tilde{\mathcal{V}}(\cdot, \mu)$ is a $C^{\frac{n}{2}+\delta}(\mathbb{R}^n, \mathbb{R})$ Hölder function for some $\delta > 0$.

- **Hypothesis C$\mathcal{V}$:** *the functional $\tilde{\mathcal{V}}$ is convex.*
- **Hypothesis QV:** *the function V, in Hypotheses $\mathcal{V}$, is radially symmetric $V(x) = \bar{V}(|x|)$, where $\bar{V}$ is a $C^1(\mathbb{R}_+, \mathbb{R})$ increasing function for which there are constants $e_1, \epsilon > 0, e_2, e_3 \geq 0$ such that:*

  i $\bar{V}(r) \geq e_1 r^{2+\epsilon} - e_2$,

  ii $\bar{V}'(r) \leq e_3 (\bar{V}(r))^{\frac{3}{2}}$, $r = |x|$.

Remark 1. The conditions (2.3) are some standard requests on the weight function V for having good properties in the Sobolev and Besov spaces on $\mathbb{R}^n$ with weight V (see, i.e., [67; 68; 69]). We use some of these properties in an essential way in Lemma 36 below.

Remark 2. An important consequence of Hypothesis $\mathcal{V}i$ is that if μ_k is a sequence in $\mathcal{P}(\mathbb{R}^n)$ converging weakly to μ, as $k \to +\infty$, then for any compact set $K \subset \mathbb{R}^n$ we have $\sup_{x \in K} |\mathcal{V}(x, \mu) - \mathcal{V}(x, \mu_k)| \to 0$. This fact is a consequence of the Prokhorov theorem (which says that $\mathcal{P}(\mathbb{R}^n)$ is a complete metric space) and of the Heine-Cantor theorem (which says that a continuous function from a compact metric space to a metric space is uniformly continuous).

Remark 3. Hypothesis C$\mathcal{V}$ is essentially used in two points of the present paper: in Theorem 14, where it is exploited for proving the uniqueness of the minimizer ρ_0, and in Theorem 25, where the uniqueness proved in Theorem 14 is applied to prove that potentials of the form (2.7) (below) satisfy the value functions convergence condition (5.1). In both cases Hypothesis C$\mathcal{V}$ guarantees uniqueness of the minimizer in the limit problem. If we do not assume Hypothesis C$\mathcal{V}$ we have to consider *relaxed controls* (see [7] for the Markovian ergodic case and [44] for controlled McKean-Vlasov dynamics).

It is important also to note that a monotonicity condition is required in the mean-field games literature in order to have uniqueness of Nash equilibrium (see, e.g. [23; 16]). More precisely if $\tilde{\mathcal{V}}$ is convex then $\partial_\mu \tilde{\mathcal{V}}$ is monotone, i.e.

$$\int_{\mathbb{R}^n} [\partial_\mu \tilde{\mathcal{V}}(y, \mu) - \partial_\mu \tilde{\mathcal{V}}(y, \mu')](\mu(dy) - \mu'(dy)) \geq 0,$$

for any probability measures $\mu, \mu' \in \mathcal{P}(\mathbb{R}^n)$.

We consider the ergodic control problem given by the cost functions (1.2). Since the cost functional $J(\alpha, x_0)$ (on the left hand side of (1.2)) depends on the law of the controlled diffusion X_t of the time averaged ergodic control problem, it is legitimate to look at it as a McKean-Vlasov control problem.

We define

$$\mathfrak{J} := \text{ess sup}_{x_0 \in \mathbb{R}^n} \left(\inf_{\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)} J(\alpha, x_0) \right) \tag{2.6}$$

where ess sup is the essential supremum over $x_0 \in \mathbb{R}^n$. In the ergodic case the optimal value $\mathfrak{J}$ is the analogous of the value function of the finite time optimal control problem. With an abuse of name we call $\mathfrak{J}$ the value function associated with the problem (1.1) and the cost functional (1.2).

Remark 4. There are two important observations to do about the initial conditions chosen in the definition of value function (2.6). The first one is that the function $x_0 \longmapsto \inf_{\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)} J(\alpha, x_0)$ is almost surely constant in x_0 with respect to the Lebesgue measure (see Theorem 17). This means that the ess $\sup_{x_0 \in \mathbb{R}^n}$ is used only to exclude a set of measure zero with respect to x_0.

The second observation is that, although in Section 3 we consider only deterministic initial conditions, it is possible to extend, in a straightforward way, our analysis by considering

$$\bar{J}(\alpha, p) := \limsup_{T \to +\infty} \frac{1}{T} \left(\int_0^T \mathbb{E}_{X_0 \sim p(x)dx} \left[\frac{|\alpha(X_t)|^2}{2} + \mathcal{V}(X_t, \text{Law}(X_t)) \right] dt \right),$$

where the process X_t has an initial probability law, $\text{Law}(X_0)$, which is absolutely continuous with respect to Lebesgue measure of the form $p(x)dx$, with p a positive Lebesgue integrable function on $\mathbb{R}^n$, and such that $\int_{\mathbb{R}^n} V(x)p(x)dx < +\infty$. Indeed in both Theorem 7 and Lemma 11 (below) we can replace the deterministic

initial condition with a random one, of the previous type, obtaining the corresponding statement. This fact proves that

$$\mathfrak{J} = \inf_{\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)} \bar{J}(\alpha, p),$$

for any $p \in L^1(\mathbb{R}^n)$, where $\mathfrak{J}$ is the same constant as in definition (2.6). In this paper, we decided to treat in detail only the case of deterministic initial conditions in order to simplify the treatment of the general problem.

2.1. A family of potentials satisfying Hypotheses $\mathcal{V}$, C$\mathcal{V}$ and Q$\mathcal{V}$

In this section we discuss a class of functionals $\mathcal{V}$ satisfying Hypotheses $\mathcal{V}$ and C$\mathcal{V}$. More precisely we consider the functionals $\mathcal{V}$ having the following form

$$\mathcal{V}(x, \mu) = V_0(x) + \int_{\mathbb{R}^n} v_0(y)\mu(dy) + \int_{\mathbb{R}^n} v_1(x-y)\mu(dy), \tag{2.7}$$

where $V_0, v_0, v_1 \in C^{\frac{n}{2}+\epsilon}(\mathbb{R}^n), \epsilon > 0$ and $\mu \in \mathcal{M}_c(\mathbb{R}^n)$ (where $\mathcal{M}_c(\mathbb{R}^n)$ is the space of signed measures on $\mathbb{R}^n$ having total mass less than $c \in \mathbb{R}_+$) . Furthermore we require that V_0 grows to plus infinity as $|x| \to +\infty$, and there is a function V, satisfying the relation (2.3) and Hypothesis QV, such that $V_0(x) \sim V(x)$ as $|x| \to +\infty$ (where $\sim$ stands for $V_0(x)$ is bounded from above and below by positive constants times $V(x)$ as $|x| \to +\infty$). We also assume that v_0, v_1 are bounded, $v_1(x) = v_1(-x)$ and that there exists a positive measure π on $\mathbb{R}^n$ such that, for any $x \in \mathbb{R}^n$, $v_1(x) = \int_{\mathbb{R}^n} e^{-ikx}\pi(dk)$ (i.e. v_1 is the Fourier transform of a positive measure).

Theorem 5. *The functional $\mathcal{V}$ of the form* (2.7) *under the above assumptions on V_0, v_0, v_1 satisfies Hypotheses $\mathcal{V}$ and C$\mathcal{V}$.*

Proof. Hypothesis $\mathcal{V}i$ follows from the fact that $\mathcal{V}(x, \cdot)$ is a sum of affine bounded functionals on $\mathcal{M}_c(\mathbb{R}^n)$. Since v_0, v_1 are bounded and V_0 grows at $+\infty$ when $|x| \to +\infty$, $\mathcal{V}$ satisfies Hypothesis $\mathcal{V}ii$. By an explicit computation we have

$$\partial_\mu(\mathcal{V})(x, y, \mu) = v_0(y) + v_1(x-y)$$

hence, since, for the previous assumptions, v_0 and v_1 are bounded and regular enough, $\mathcal{V}$ satisfies $\mathcal{V}iii$. Furthermore we get, by the definition (2.1) and the fact that the integral of v_0 in (2.7) is constant

$$\tilde{\mathcal{V}}(\mu) = \int_{\mathbb{R}^n} (V_0(x) + v_0(x))\mu(dx) + \int_{\mathbb{R}^{2n}} v_1(x-y)\mu(dx)\mu(dy)$$

and so

$$\partial_\mu^2(\tilde{\mathcal{V}})(x, y, \mu) = 2v_1(x-y).$$

The previous equation implies that, if the bilinear form

$$B(\tilde{\mu}, \tilde{\nu}) = 2\int_{\mathbb{R}^n} v_1(x-y)\tilde{\nu}(dx)\tilde{\mu}(dy), \quad \tilde{\mu}, \tilde{\nu} \in \mathcal{M}_{2c}(\mathbb{R}^n)$$

is nonnegative definite, then $\tilde{\mathcal{V}}$ is convex having nonnegative definite second differential. Since v_1 is continuous and bounded, by Theorem XVIII of Chapter VII in [70] (see also [63, Theorem IX.10]), the form B is positive definite if and only if v_1 is a positive definite function. By Bochner's theorem (see, e.g., [63, Theorem IX.9]), v_1 is a positive definite continuous function if and only if it is the Fourier transform of a positive measure. This complete the proof of the theorem. □

Remark 6. Using Theorem 5 it is possible to build other functionals satisfying Hypotheses $\mathcal{V}$, C$\mathcal{V}$ and QV. Indeed we can, e.g., compose functionals of the form (2.7) with the derivatives of an homogeneous symmetric polynomial $P : \mathbb{R}^k \to \mathbb{R}$ which is convex and it has positive partial derivatives on $\mathbb{R}_+^k$. More precisely let $v_1, ..., v_k$ be positive functions satisfying the same conditions of v_1 in Theorem 5, and consider

$$\mathcal{V}(x, \mu) = V_0(x) +$$
$$+ \sum_{\ell=1}^k \partial_{z_\ell}(P_k)\left(\int_{\mathbb{R}^{2n}} v_1(y_1' - y_1)\mu(dy_1')\mu(dy_1), \cdots, \int v_k(y_k' - y_k)\mu(dy_k')\mu(dy_k)\right)\int_{\mathbb{R}^n} v_\ell(x - y_\ell)\mu(dy_\ell). \tag{2.8}$$

Using the fact that P_k is an homogeneous polynomial, we get that the functional $\tilde{\mathcal{V}}$ associated with the operator $\mathcal{V}$ given in (2.8), we get

$$\tilde{\mathcal{V}}(\mu) = \int_{\mathbb{R}^k} V_0(y)\mu(dy) + \deg(P_k)P_k\left(\int_{\mathbb{R}^{2n}} v_1(y_1' - y_1)\mu(dy_1')\mu(dy_1), \cdots, \int v_k(y_k' - y_k)\mu(dy_k')\mu(dy_k)\right).$$

This implies that $\tilde{\mathcal{V}}$ is convex since it is the sum of a linear function, and the composition of positive convex functionals $\int_{\mathbb{R}^{2n}} v_\ell(y_\ell - y'_\ell)\mu(dy_\ell)\mu(dy'_\ell)$ and a convex function with positive derivatives $P_k : \mathbb{R}_+^k \to \mathbb{R}$. Since the linear combination (with positive coefficients) of convex functional is convex we can use the previous construction to build general non quadratic, and also non polynomial, cost functionals.

3. The McKean-Vlasov optimal control problem

3.1. The ergodic control problem

We are searching for the control function $\alpha \in C^1$ which minimizes the functional (1.2). First of all we need some results and notations concerning equations of the form (1.1) when $\alpha \in C^1$ is admitting an invariant measure. We denote by μ_{t,x_0} the probability measure on $\mathbb{R}^n$ giving the distribution of X_t when $X_0 = x_0$, $x_0 \in \mathbb{R}^n$. We also write $\tilde{\mu}_{t,x_0}$ for the following time averaged measure

$$\tilde{\mu}_{t,x_0}(x) = \frac{1}{t} \int_0^t \mu_{\tau,x_0}(dx)d\tau, x \in \mathbb{R}^n \text{ and } t \in \mathbb{R}_+$$

We denote by T_t, $t \in \mathbb{R}_+$, the (sub)Markovian semigroup associated with SDE (1.1), namely if $f \in L^1(\mathbb{R}^n, \mu_{t,x}) \equiv L^1(\mu_{t,x})$ we have

$$T_t(f)(x) = \int_{\mathbb{R}^n} f(y)\mu_{t,x}(dy) = \mathbb{E}_x[f(X_t)]. \tag{3.1}$$

We denote by $L_\alpha : C^\infty(\mathbb{R}^n) \to C^1(\mathbb{R}^n)$ the operator

$$L_\alpha(g)(x) = \Delta g(x) + \sum_{i=1}^n \alpha^i(x)\partial_{x_i}(g)(x), \quad g \in C^\infty(\mathbb{R}^n) \tag{3.2}$$

We say that a probability measure μ is an invariant measure for the process X_t, $t \geq 0$, or equivalently, for the semigroup T_t if, for any bounded measurable function f, we have

$$\int_{\mathbb{R}^n} T_t(f)(x)\mu(dx) = \int_{\mathbb{R}^n} f(x)\mu(dx).$$

We say that a probability measure μ is infinitesimal invariant for T_t, $t \geq 0$, and we write $L_\alpha^*(\mu) = 0$, if for any $g \in C^\infty(\mathbb{R}^n)$ with compact support we have

$$\int_{\mathbb{R}^n} L_\alpha(g)(x)\mu(dx) = 0. \tag{3.3}$$

Proposition 7. *Consider an SDE of the form* (1.1) *with* $\alpha \in C^1$, *and suppose that it admits an invariant measure* μ. *Then the following assertions hold:*

  i T_t *is strong Feller,*
  ii μ *is the unique ergodic invariant measure of* T_t,
  iii μ *is absolutely continuous with respect to the Lebesgue measure,*
  iv *for any* $x_0 \in \mathbb{R}^n$, $\tilde{\mu}_{t,x_0} \to \mu$ *weakly as* $t \to +\infty$,
  v *for any* $x_0 \in \mathbb{R}^n$, $\mu_{t,x_0} \to \mu$ *weakly as* $t \to +\infty$,
  vi *if further* $|\alpha|^2 \in L^1(\mu)$ *then for any* $f \in L^1(\mu)$ *we have* $\lim_{t\to+\infty} \frac{1}{t} \int_0^t T_s(f)(x_0) = \int_{\mathbb{R}^n} f(x)\mu(dx)$ *for* μ-*almost all* $x_0 \in \mathbb{R}^n$,
  vii *finally* μ *is the unique invariant measure of* T_t *if and only if it is the unique solution to the equation* $L_\alpha^*(\mu) = 0$.

Proof. By [55, Proposition 2.2.12] T_t is irreducible and strong Feller. This implies that X_t has an unique ergodic invariant measure, from Doob's Theorem in [30, Theorem 4.2.1], which means that μ is the unique solution to the Fokker-Planck equation $L_\alpha^*(\mu) = 0$, where L_α is the infinitesimal generator of T_t (which is the unique extension of the operator (3.2)) and L_α^* its adjoint, (proving the point *ii*). Furthermore, since α is C^1 and L_α is uniformly elliptic, by [14, Corollary 1.5.3], we have that μ is absolutely continuous with respect to Lebesgue measure. Points *iv* and *v* are consequences of [30, Theorem 4.2.1]. Furthermore, using the fact that $|\alpha|^2 \in L^1(\mu)$ and by Theorem 5.2.9 of [14], the semigroup T_t is a strongly continuous semigroup on $L^1(\mu)$. By Remark 1 in [76, Chapter XII Section 1], this implies point *vi*. The point *vii* follows by the uniqueness of the invariant measure. $\square$

Remark 8. By a classical result a sufficient condition for the existence of an invariant measure is that α is of the form $\alpha = -DU - G$, with $U \in C_1^{1+\beta}(\mathbb{R}^N)$ for some $\beta \in (0,1)$, $G \in C^1(\mathbb{R}^N, (\mathbb{R}^N)$, $e^{-U} \in L^1(\mathbb{R}^N)$, $|G|e^{-U} \in L^1(\mathbb{R}^N)$, $\operatorname{div} G =< G, DU >$. In this case $\mu(dy) = \frac{\exp(-U(x)dy)}{\int \exp -U(x)dx}$ is symmetric in $L^2(\mu)$ (see, e.g., [55], Chapter 8 Theorem 8.1.26).

Remark 9. It is important to note that the condition $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ is an essential hypothesis in Proposition 7 (and consequently in Lemma 11 and Theorem 17 below). Indeed, when α is in general only measurable and L^2 with respect (some) invariant measure μ there is the possibility of multiple invariant measures even in one dimension (see, e.g., the discussions in [14, Chapter 4]).

Remark 10. Since, when $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, and by Proposition 7 *vii*, the two notions of invariant measures, namely then ones given by equations (3.1) and (3.3) respectively, are equivalent, hereafter we call invariant measure a measure which respect both equations (3.1) and (3.3).

In the next lemma we shall provide a sufficient condition for the existence of an invariant measure for the SDE (1.1) admitting a probability density. This allows to obtain a cost functional expressed in terms of a probability density notably simplifying our minimization problem. Hereafter we use the following abuse of notation: If μ is a measure on $\mathbb{R}^n$ absolutely continuous with respect to the Lebesgue measure with density ρ (namely $d\mu(x) = \rho(x)dx$) we write $\mathcal{V}(x, \rho)$ instead of the more precise $\mathcal{V}(x, \mu) = \mathcal{V}(x, \rho dx)$.

Lemma 11. *Under hypotheses $\mathcal{V}i$ and $\mathcal{V}ii$, if $J(\alpha, x_0)$ as given by (1.2) (with $\alpha \in C^1$) is not equal to $+\infty$ there exists an unique and ergodic invariant probability density measure $\rho_\alpha \in W^{1,\frac{n}{2}}(\mathbb{R}^n)$ for the SDE (1.1) so that $\mu_\alpha(dx) = \rho_\alpha(x)dx$, with μ_α the invariant ergodic probability measure for the SDE (1.1). Furthermore we have*

$$\tilde{J}(\alpha, \rho_\alpha) \leq J(\alpha, x_0)$$

for almost all $x_0 \in \mathbb{R}^n$ with respect to Lebesgue measure, where

$$\tilde{J}(\alpha, \rho_\alpha) := \int_{\mathbb{R}^n} \left(\frac{|\alpha(x)|^2}{2} + \mathcal{V}(x, \rho_\alpha) \right) \rho_\alpha(x)dx. \tag{3.4}$$

Proof. Under hypothesis $\mathcal{V}ii$ when $J(\alpha, x_0)$ is finite, for any $x_0 \in \mathbb{R}^n$, $\tilde{\mu}_{t,x_0}$, indexed by $t \in \mathbb{R}_+$, is a family of tight measures. Indeed we have that, for any $t > 0$:

$$\begin{aligned}
\int_{\mathbb{R}^n} V(x)\tilde{\mu}_{t,x_0}(dx) =& \frac{1}{t} \int_0^t \int_{\mathbb{R}^n} V(x)\mu_{\tau,x_0}(dx)d\tau \\
\leq& \frac{1}{t} \int_{\mathbb{R}^n} \int_0^t \mathcal{V}(x, \mu_{\tau,x_0})\mu_{\tau,x_0}(dx)d\tau + c_1 \\
\leq& \frac{1}{t} \left(\int_0^t \mathbb{E}_{x_0} \left[\frac{|\alpha(X_\tau)|^2}{2} + \mathcal{V}(X_\tau, \operatorname{Law}(X_\tau)) \right] d\tau \right) + c_1 < C
\end{aligned}$$

for some $C \in \mathbb{R}$, where in the last step we used that $J(\alpha, x_0) < +\infty$.

Since V is a function growing to infinity as $|x| \to +\infty$, the family $(\tilde{\mu}_{t,x_0}(dx), t > 0)$ is necessarily tight. Now let $\mu(dx)$ be any weak limit of a subsequence of $\tilde{\mu}_{t,x_0}(dx)$, as $t \to +\infty$, then $\mu(dx)$ is an invariant probability measure for equation (1.1). Indeed let f be a C^∞ function with compact support then, by Itô formula, for $t > 0$:

$$\int_{\mathbb{R}^n} L_\alpha(f)(x)\tilde{\mu}_{t,x_0}(dx) = \frac{1}{t}(\mathbb{E}[f(X_t)] - f(x_0)).$$

Since f has compact support we have that $\lim_{t \to +\infty} \frac{1}{t}(\mathbb{E}[f(X_t)] - f(x_0)) = 0$, which implies, α being locally bounded, that

$$0 = \lim_{t \to +\infty} \int_{\mathbb{R}^n} L_\alpha(f)(x)\tilde{\mu}_{t,x_0}(dx) = \int_{\mathbb{R}^n} L_\alpha(f)(x)\mu(dx).$$

This means that $L_\alpha^*(\mu) = 0$ and thus μ is an invariant probability measure for equation (1.1). By Proposition 7 *iii*, there exists a positive $L^1(\mathbb{R}^n)$ function $\rho_\alpha(x)$ such that $\mu(dx) = \rho_\alpha(x)dx$.

What remains to be proved is that $J(\alpha, x_0) \geq \tilde{J}(\alpha, \rho_\alpha)$ (Lebesgue-)almost surely with respect to $x_0 \in \mathbb{R}^n$.

We have that $\liminf_{t \to +\infty} \int_{\mathbb{R}^n} \frac{|\alpha(x)|^2}{2} \tilde{\mu}_{t,x_0}(dx) \geq \int_{\mathbb{R}^n} \frac{|\alpha(x)|^2}{2}\rho_\alpha(x)dx$. Indeed, for any $N \in \mathbb{N}$:

$$\begin{aligned}
\int_{\mathbb{R}^n} \frac{|\alpha(x)|^2 \wedge N}{2}\rho_\alpha(x)dx &= \lim_{t \to +\infty} \int_{\mathbb{R}^n} \frac{|\alpha(x)|^2 \wedge N}{2} \tilde{\mu}_{t,x_0}(dx) \\
&\leq \liminf_{t \to +\infty} \int_{\mathbb{R}^n} \frac{|\alpha(x)|^2}{2} \tilde{\mu}_{t,x_0}(dx) < +\infty.
\end{aligned}$$

Since $\lim_{N\to+\infty}\int_{\mathbb{R}^n}\frac{|\alpha(x)|^2\wedge N}{2}\rho_\alpha(x)dx=\int_{\mathbb{R}^n}\frac{|\alpha(x)|^2}{2}\rho_\alpha(x)dx$ the stated inequality is proved. Now we want to prove that

$$\limsup_{t\to+\infty}\int_{\mathbb{R}^n}\mathcal{V}(x,\mu_{t,x_0})\mu_{t,x_0}(dx)=\int_{\mathbb{R}^n}\mathcal{V}(x,\mu)\rho_\alpha(x)dx, \tag{3.5}$$

almost surely with respect to $x_0\in\mathbb{R}^n$. Let $t_m\to+\infty$ be a sequence in $\mathbb{R}_+$ which realizes the $\limsup$ in (3.5). By Proposition 7 *vi* and denoting by B_K the ball of radius $K\in\mathbb{N}$ and center in 0 we have that

$$\lim_{m\to+\infty}\int_{\mathbb{R}^n\setminus B_K}V(x)\tilde{\mu}_{t_m,x_0}(dx)=\lim_{m\to+\infty}\frac{1}{t_m}T_{t_m}(\mathbb{I}_{\mathbb{R}^n\setminus B_K}V)(x_0)$$

$$=\int_{\mathbb{R}^n\setminus B_K}V(x)\rho_\alpha(x)dx,$$

for almost all $x_0\in\mathbb{R}^n$ and for all $K\in\mathbb{N}$. Since the positive measure $V(x)\rho_\alpha(x)dx$ is regular, this means that the sequence of positive measures $V(x)\tilde{\mu}_{t_m,x_0}(dx)$ is tight. By Hypothesis $\mathcal{V}ii$, the tightness of $V(x)\tilde{\mu}_{t_m,x_0}(dx)$ implies the tightness of the sequence of signed measures (with total mass uniformly bounded) $\frac{1}{t_m}\int_0^{t_m}\mathcal{V}(x,\mu_{s,x_0})\mu_{s,x_0}(dx)ds$. On the other hand, by Remark 2 (in Section 2) and using the fact that, by Proposition 7 *iv*, $\mu_{t,x_0}\to\mu$ weakly, as $t\to+\infty$, we have

$$\lim_{m\to+\infty}\frac{1}{t_m}\int_0^{t_m}\int_{\mathfrak{K}}\mathcal{V}(x,\mu_{s,x_0})\mu_{s,x_0}(dx)ds=$$

$$=\lim_{m\to+\infty}\left(\int_{\mathfrak{K}}\mathcal{V}(x,\mu)\tilde{\mu}_{s,x_0}(dx)ds+\frac{1}{t_m}\int_0^{t_m}\int_{\mathfrak{K}}(\mathcal{V}(x,\mu_{s,x_0})-\mathcal{V}(x,\mu))\mu_{s,x_0}(dx)ds\right)$$

$$=\int_{\mathfrak{K}}\mathcal{V}(x,\mu)\rho_\alpha(x)dx \quad (3.6)$$

for any compact set $\mathfrak{K}\subset\mathbb{R}^n$. Since $\frac{1}{t_m}\int_0^{t_m}\mathcal{V}(x,\mu_{s,x_0})\mu_{s,x_0}(dx)ds$ has a uniformly bounded mass and is tight, relation (3.6) implies that $\frac{1}{t_m}\int_0^{t_m}\mathcal{V}(x,\mu_{s,x_0})\mu_{s,x_0}(dx)ds$ converges as $m\to\infty$ to $\mathcal{V}(x,\mu)\rho_\alpha(x)(dx)$, weakly for (Lebesgue) almost all $x_0\in\mathbb{R}^n$. This proves equality (3.5) and concludes the proof. $\square$

3.2. A lower bound for the functional (3.4)

In order to minimize the functional (3.4) with respect to ρ, for $\rho\in W^{1,\frac{n}{2}}(\mathbb{R}^n)$, $\rho(x)\geq 0$ and $\int_{\mathbb{R}^n}\rho(x)dx=1$, we set

$$\mathcal{C}_\rho=\{\alpha\in C^1(\mathbb{R}^n,\mathbb{R}^n),\ L_\alpha^*(\rho)=0\text{ and }|\tilde{J}(\alpha,\rho)|<+\infty\}. \tag{3.7}$$

Then $\mathcal{C}_\rho$ is the subset of $C^1(\mathbb{R}^n,\mathbb{R}^n)$ vector fields $\alpha_\rho\in\mathcal{C}_\rho$ such that $L_{\alpha_\rho}^*(\rho)=0$ (where $L_{\alpha_\rho}^*$ is the adjoint of the infinitesimal generator of the equation (1.1) and the previous equality is understood in a distributional sense) and $|\tilde{J}(\alpha_\rho,\rho)|<+\infty$.

Remark 12. Suppose that $\alpha\in C^1(\mathbb{R}^n,\mathbb{R}^n)$ such that $J(\alpha,x_0)<+\infty$, then by Lemma 11 there is a unique positive probability density ρ_α which is invariant and thus, since $\alpha\in C^1(\mathbb{R}^n,\mathbb{R}^n)$ by Proposition 7, it satisfies the equation $L_\alpha^*(\rho_\alpha)=0$. This implies that $\alpha\in\mathcal{C}_{\rho_\alpha}$, where $\mathcal{C}_{\rho_\alpha}$ is defined by equation (3.7) with $\rho=\rho_\alpha$.

We introduce the following energy functional, for $\rho\in W^{1,\frac{n}{2}}(\mathbb{R}^n)$,

$$\mathcal{E}(\rho):=\mathcal{E}_K(\rho)+\mathcal{E}_P(\rho)=\int_{\mathbb{R}^n}\frac{|\nabla\rho|^2}{\rho}dx+\int_{\mathbb{R}^n}\mathcal{V}(x,\rho)\rho(x)dx. \tag{3.8}$$

where the two terms on the right hand side correspond to the kinetic $\mathcal{E}_K(\rho)$ and potential $\mathcal{E}_P(\rho)$ energies, respectively.

The next lemma states a useful monotonicity property of the cost functional $\tilde{J}$.

Lemma 13. *For any given* $\rho\in W^{1,\frac{n}{2}}(\mathbb{R}^n)$ *we have*

$$\mathcal{E}(\rho)=\tilde{J}\left(\frac{\nabla\rho}{\rho},\rho\right)\leq\inf_{\alpha\in\mathcal{C}_\rho}\tilde{J}(\alpha,\rho),$$

where $\tilde{J}(\alpha,\rho)$ *is defined in* (3.4).

Proof. By [14, Chapter 3, Theorem 3.1.2], if ρ is the density of the invariant measure of the SDE (1.1) we have that

$$\int_{\mathbb{R}^n} \frac{|\nabla \rho(x)|^2}{\rho^2(x)} \rho(x) dx \leq \int_{\mathbb{R}^n} |\alpha(x)|^2 \rho(x) dx,$$

for any $\alpha \in \mathcal{C}_\rho$, with the equality holding if and only if $\alpha = \frac{\nabla \rho}{\rho}$. Since $\int_{\mathbb{R}^n} \mathcal{V}(x, \mu)\rho(x)dx$ does not depend on α but only on the invariant measure $\rho(x)dx$, the theorem is proved. $\square$

3.3. Minimizer of the energy functional

We want to minimize the function $\mathcal{E}(\rho)$ given by (3.8) under the condition $\int_{\mathbb{R}^n} \rho(x)dx = 1$. It is useful to introduce the following variable $\phi = \sqrt{\rho}$. With this notation the energy functional (3.8) becomes

$$\mathcal{E}(\phi^2) = \int_{\mathbb{R}^n} \left(\frac{|\nabla \phi|^2}{2} + \mathcal{V}(x, \phi^2)\phi^2(x) \right) dx, \tag{3.9}$$

with $\phi \in L^2(\mathbb{R}^n)$ satisfying the condition $\int_{\mathbb{R}^n} \phi^2(x)dx = 1$.

The following result states that the above energy functional admits a unique minimizer which is strictly positive.

Lemma 14. *Under hypotheses $\mathcal{V}$ and $C\mathcal{V}$ the variational problem (3.8), with $\phi \in L^2(\mathbb{R}^n)$ satisfying the condition $\int_{\mathbb{R}^n} \phi^2(x)dx = 1$, admits a unique minimizer $\rho_0 = \phi_0^2$. Furthermore ϕ_0 is $C^{2+\epsilon}(\mathbb{R}^n)$ for some $\epsilon > 0$, it is strictly positive and satisfies (weakly) the equation*

$$-\Delta\phi_0(x) + 2\mathcal{V}(x, \phi_0^2)\phi_0(x) + 2\int_{\mathbb{R}^n} \partial_\mu \mathcal{V}(y, x, \phi_0^2)\phi_0^2(y)dy\phi_0(x) = \mu_0\phi_0(x), \tag{3.10}$$

where the uniquely determined constant μ_0 given by

$$\mu_0 = 2\mathcal{E}(\phi_0^2) + \int_{\mathbb{R}^n} \partial_\mu \mathcal{V}(y, x, \phi_0^2)\phi_0^2(y)\phi_0^2(x)dydx \tag{3.11}$$

Proof. By Hypothesis $C\mathcal{V}$ the functional $\mathcal{E}_P(\rho)$ is convex and by the property of Fisher information (see Theorem 31 below), $\mathcal{E}_K(\rho)$ is convex and strictly convex when it is finite. Furthermore by Hypothesis $\mathcal{V}ii$ $\mathcal{E}$ is coercive in ϕ_0 (in the sense that $\mathcal{E}(\phi^2) \geq C||\phi^2||_{H^1}$). This implies that there exists a unique minimizer $\phi_0 = \sqrt{\rho_0}$.

On the other hand, making a variation of the form $\phi_0 + \epsilon\delta\phi$, where $\epsilon > 0$ and $\delta\phi$ is supposed to be a smooth compactly supported function, under the additional constraint given by the normalization condition for $(\phi_0)^2$, by the regularity property given by Hypothesis $\mathcal{V}iii$, the minimizer ϕ_0 must satisfy (in a weak sense) equation (3.10). For determining the Lagrange multiplier μ_0 it is sufficient to multiply both sides of equation (3.10) by ϕ_0 and then integrate by parts.

Using a bootstrap argument, beginning by $(\partial_\mu\tilde{\mathcal{V}})(\cdot, \phi_0^2) \in C^{\frac{n}{2}+\epsilon'}$, for some $\epsilon' > 0$ by Hypothesis $\mathcal{V}iii$ and by elliptic regularization property of the Laplacian (see Theorem 8.10 in [37]), we obtain that $\phi_0 \in H_{loc}^{\frac{n}{2}+\epsilon}(\mathbb{R}^2)$ and thus $\phi_0 \in C^\epsilon(\mathbb{R}^n)$. Exploiting the regularity results for the Poisson equation (see Theorem 4.3 in [37]), we have that $\phi_0 \in C^{2+\epsilon}(\mathbb{R}^n)$.

Finally, equation (3.10) implies that ϕ_0 is the ground state of a quantum mechanical system on $\mathbb{R}^n$ with potential $2\partial_\mu\tilde{\mathcal{V}}(x, \phi_0^2)$ (where $\tilde{\mathcal{V}}$ is defined in (2.1)). Since, by Hypotheses $\mathcal{V}ii$ and $\mathcal{V}iii$, $2\mathcal{V}(x, \phi_0^2) + 2\partial_\mu\tilde{\mathcal{V}}(x, \phi_0^2)$ is bounded from below and diverges to infinity as $|x| \to +\infty$, by [64, Theorem XIII.47] we have that ϕ_0 is strictly positive. $\square$

Remark 15. In Lemma 14 Hypothesis $C\mathcal{V}$ is only used to prove the uniqueness of the minimizer ρ_0. Indeed in order to prove existence and positivity of ϕ_0 we need only Hypotheses $\mathcal{V}$.

Remark 16. The minimizer ρ_0 in Lemma 14 satisfies the following equation

$$-\Delta\rho_0(x) + \frac{|\nabla\rho_0(x)|^2}{2\rho_0(x)} + \mathcal{V}(x, \rho_0)\rho_0(x) + \int_{\mathbb{R}^n} \partial_\mu\mathcal{V}(y, x, \rho_0)\rho_0(y)dy\rho_0(x) = \mu_0\rho_0(x), \tag{3.12}$$

as easily deduced from (3.10).

3.4. Existence and uniqueness of the optimal control

Finally we obtain the explicit form of the optimal control:

Theorem 17. *Under Hypotheses $\mathcal{V}$ and $\mathcal{CV}$, the logarithmic gradient of the unique minimizer $\rho_0 = \phi_0^2$ of $\mathcal{E}$, that is $\alpha = \frac{\nabla \rho_0}{\rho_0}$, is the optimal control for the problem (1.2) for almost every $x_0 \in \mathbb{R}^n$ with respect to the Lebesgue measure.*

In order to prove Theorem 17 we need the following lemma.

Lemma 18. *Under Hypotheses $\mathcal{V}$ and $\mathcal{CV}$ we have that*

$$J\left(\frac{\nabla \rho_0}{\rho_0}, x_0\right) = \mathcal{E}(\rho_0), \tag{3.13}$$

where $\rho_0 = \phi_0^2$ is the unique minimizer of $\mathcal{E}$.

Proof. We have that $\mu(dx) = \rho_0(x)dx$ is the unique ergodic invariant probability measure of the strong Feller SDE (1.1) with $\alpha = \frac{\nabla \rho_0}{\rho_0}$. By the definition of $\mathcal{E}$ and equation (3.10) we have that $x \longmapsto |\alpha(x)|^2 = \frac{|\nabla \rho_0(x)|^2}{\rho_0(x)^2}$ belongs to $L^1(\mu)$. This implies, using Proposition 7 *vi*, that we have

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t \mathbb{E}_{x_0}[|\alpha(X_s)|^2]ds = \lim_{t \to +\infty} \frac{1}{t} \int_0^t T_s(|\alpha(x)|^2)(x_0)ds = \int_{\mathbb{R}^n} |\alpha(x)|^2 \rho_0(x)dx,$$

for (Lebesgue) almost every $x_0 \in \mathbb{R}^n$ (this is due to the fact that μ is absolutely continuous and ρ_0 is strictly positive). The proof of the fact that

$$\limsup_{t \to +\infty} \frac{1}{t} \int_0^t \mathbb{E}_{x_0}[\mathcal{V}(X_s, \mathrm{Law}(X_s)]ds = \int_{\mathbb{R}^n} \mathcal{V}(x, \mu_0)\rho_0(x)dx,$$

is given in Lemma 11 (see equation (3.5) and what follows from it). $\qquad \square$

Proof of Theorem 17. By Lemma 18, and the definition of $\mathfrak{J}$ (given in equation (2.6)) we have that

$$\mathfrak{J} \leq \mathrm{ess}\ \sup_{x_0 \in \mathbb{R}^n} J\left(\frac{\nabla \rho_0}{\rho_0}, x_0\right) = \mathcal{E}(\rho_0). \tag{3.14}$$

In order to prove the statement of the theorem, it is sufficient to prove that

$$\mathcal{E}(\rho_0) \leq \mathfrak{J},$$

indeed, by Lemma 18 and inequality (3.14), this implies that $\mathfrak{J} \leq \mathrm{ess}\ \sup_{x_0 \in \mathbb{R}^n} J\left(\frac{\nabla \rho_0}{\rho_0}, x_0\right)$ and thus the thesis. By Lemma 11, we have $\tilde{J}(\alpha, \rho_\alpha) \leq J(\alpha, x_0)$ and by Lemma 13, and since, by Remark 12, $\alpha \in \mathcal{C}_{\rho_\alpha}$, we get, for any fixed $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ such that $J(\alpha, x_0) < +\infty$,

$$\mathcal{E}(\rho_\alpha) = \tilde{J}\left(\frac{\nabla \rho_\alpha}{\rho_\alpha}, \rho_\alpha\right) \leq \inf_{\hat{\alpha} \in \mathcal{C}_{\rho_\alpha}} \tilde{J}(\hat{\alpha}, \rho) \leq \tilde{J}(\alpha, \rho_\alpha).$$

Combining the previous two inequalities and Lemma 14, we obtain that, for any $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ such that $J(\alpha, x_0) < +\infty$,

$$\mathcal{E}(\rho_0) \leq \mathcal{E}(\rho_\alpha) \leq \tilde{J}(\alpha, \rho_\alpha) \leq \mathrm{ess}\ \sup_{x_0 \in \mathbb{R}^n} J(\alpha, x_0).$$

Taking the inf over $\alpha \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ from the previous inequality we get $\mathcal{E}(\rho_0) \leq \mathfrak{J}$. $\qquad \square$

Remark 19. An important consequence of Theorem (17) is that under Hypotheses $\mathcal{V}$ and $\mathcal{CV}$ we have that

$$\mathfrak{J} = \mathcal{E}(\rho_0) = \inf_{\phi \in H^1(\mathbb{R}^n), \int \phi^2 dx = 1} \mathcal{E}(\phi^2),$$

where $\mathfrak{J}$ is the value function associated with the problem (1.1) and the cost functional (1.2), defined by (2.6).

4. The N-particles approximation

In order to rigorously justify the limit McKean-Vlasov optimal control problem discussed in Section 3 , in this section we propose for it a natural many particles approximation. We consider the process $X_t = (X_t^1, ..., X_t^N) \in \mathbb{R}^{nN}$ satisfying the SDE

$$dX_t^i = A_N^i(X_t)dt + \sqrt{2}dW_t^i, \tag{4.1}$$

where $A_N = (A_N^1, ..., A_N^N) : \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ is a $C^{1+\epsilon}$ function, for some $\epsilon > 0$, and the $W_t^i, i = 1, \dots, N$ are independent Brownian motions taking values in $\mathbb{R}^n$.

Remark 20. It is important to note that, although the Brownian motions $W_t^i, i = 1, \ldots, N$ are independent, the processes $X_t^i, i = 1, \ldots, N$ are in general not independent since we do not require that A_N^i depends only on the variable x^i, but it can in general depend on all the variables $(x^1, \ldots, x^n)$.

If $\mathcal{V}$ is a functional satisfying Hypotheses $\mathcal{V}$, we introduce the functions sequence

$$\mathcal{V}_N(x) = \sum_{i=1}^N \mathcal{V}\left(x_i, \frac{1}{N-1}\sum_{k=1,k\neq i}^N \delta_{x^i}\right),$$

where $x = (x^1, \ldots, x^N) \in \mathbb{R}^{nN}$, $N \geq 2$.

We consider the (normalized with respect to the number of particles N) ergodic control problem

$$J_N(A_N, x_0) = \limsup_{T \to +\infty} \frac{1}{NT}\int_0^T \mathbb{E}_{x_0}\left[\frac{|A_N(X_t)|^2}{2} + \mathcal{V}_N(X_t)\right] dt, \tag{4.2}$$

and also the (normalized) energy functional

$$\mathcal{E}_N(\rho_N) = \mathcal{E}_{K,N}(\rho_N) + \mathcal{E}_{P,N}(\rho_N) = \frac{1}{N}\left(\int_{\mathbb{R}^{nN}} \frac{|\nabla\rho_N|^2}{2\rho_N}dx + \int_{\mathbb{R}^{nN}} \mathcal{V}_N(x)\rho_N(x)dx\right), \tag{4.3}$$

where ρ_N is a positive Lebesgue integrable function such that $\int_{\mathbb{R}^{nN}} \rho_N(x)dx = 1$. We also consider the value function

$$\mathfrak{J}_N = \text{ess sup}_{x_0 \in \mathbb{R}^n}\left(\inf_{A_N \in C^1(\mathbb{R}^{nN}, \mathbb{R}^{nN})} J_N(A_N, x_0)\right). \tag{4.4}$$

Let us introduce the notation $\rho_N^{(1)}(x^1) = \int_{\mathbb{R}^{n(N-1)}} \rho_N(x^1, x^2, \ldots, x^N)dx^2\ldots dx^N$ for the one-particle probability density and let us finally put $\phi_N = \sqrt{\rho_N}$.

The next theorem, which is the analogue of Lemma 14 for our N-particles control problem, gives important properties of the minimizer of the above energy functional. In particular, since the unique minimizer is symmetric, our N-particles control problem is intrinsically symmetric: for every fixed N the diffusion components are not independent but they are identically distributed (see [58]).

Lemma 21. *Under Hypotheses $\mathcal{V}$, there exists a unique minimizer $\rho_{0,N} = \phi_{0,N}^2$ of the functional $\mathcal{E}_N$. This minimizer is symmetric in $x^1, \ldots, x^N$, it is $C^{2+\epsilon}(\mathbb{R}^{nN})$, for some $\epsilon > 0$, and it is strictly positive. Furthermore it is the only weak solution of the following linear PDE*

$$-\Delta\phi_{0,N}(x) + 2\mathcal{V}_N(x)\phi_{0,N} = \mu_N\phi_{0,N}, \tag{4.5}$$

where

$$\mu_N = 2\mathcal{E}_N(\phi_{0,N}^2).$$

Proof. Lemma 21 can be seen as a special version of Lemma 14 when $\mathcal{V}$ does not depend on ρ. The uniqueness of the minimizer is guaranteed by the fact that $\mathcal{E}_N(\phi^2)$ is quadratic with coefficients bounded from below (see, e.g., [51], Chapter 11). $\qquad\square$

Remark 22. It is important to note that, by uniqueness of the minimizer $\rho_{0,N}$ of the functional $\mathcal{E}_N$, it follows that $\rho_{0,N}$ must be invariant with respect to coordinates permutations. Indeed it is simple to prove, using convexity of the Fisher information (see below), that if $\rho_{0,N}$ is a minimizer also its symmetrization is a minimizer (see, e.g., [51], Chapter 7). An important consequence of the symmetry of $\rho_{0,N}$ with respect to coordinates permutations is that, when we consider $A_N = \frac{\nabla\rho_{0,N}}{\rho_{0,N}}$ and we start from a symmetric probability measure $p_N(x)$, as, for example, $\rho_{0,N} = p_N$ itself, then the process $(X_t^1, \ldots, X_t^N)$ is *symmetric with respect to permutations (or equivalently exchangeable)*. This observation plays a very important role in the rest of the paper.

Finally the analogue of Theorem 17 provides the optimal control.

Theorem 23. *Under Hypotheses $\mathcal{V}$, the logarithmic gradient of the unique minimizer $\rho_{0,N} = \phi_{0,N}^2$ of $\mathcal{E}_N$, that is*

$$(A_N^1, \ldots, A_N^N) = \left(\frac{\nabla_1\rho_{0,N}}{\rho_{0,N}}, \ldots, \frac{\nabla_N\rho_{0,N}}{\rho_{0,N}}\right),$$

is the optimal control of the problem (4.2).

Proof. Theorem 23 can be seen as a special version of Theorem 17 when $\mathcal{V}$ does not depend on ρ. $\qquad\square$

Remark 24. A very useful consequence of Theorem 23 is that

$$\mathfrak{J}_N := \mathcal{E}_N(\rho_{0,N}) = \inf_{\phi_N \in H^1(\mathbb{R}^{nN}), \int \phi_N^2 dx = 1} \mathcal{E}_N(\phi_N^2).$$

5. The convergence of value functions

In this section we prove the following convergence theorem.

Theorem 25. *Suppose $\mathcal{V}$ satisfies Hypotheses $\mathcal{V}$ and $C\mathcal{V}$ then we have*

$$\lim_{N\to\infty} \mathfrak{J}_N = \mathfrak{J}, \tag{5.1}$$

where $\mathfrak{J}$ is as in (2.6). Furthermore we have

$$\lim_{N\to\infty} \mathcal{E}_{K,N}(\phi_{0,N}^2) = \mathcal{E}_K(\phi_0^2) \tag{5.2}$$

$$\lim_{N\to\infty} \rho_{0,N}^{(1)}(\cdot) = \lim_{N\to\infty} \int_{\mathbb{R}^{n(N-1)}} \rho_{0,N}(\cdot, x^2, ..., x^N) dx^2 ... dx^N = \rho_0(\cdot) \tag{5.3}$$

where the last limit is understood weakly in $L^1(\mathbb{R}^n, V(x)dx)$ and ρ_0 and $\rho_{0,N}$ are the unique minimizer of the functionals $\mathcal{E}$ and $\mathcal{E}_N$ respectively (with $\mathcal{E}$ given by (3.9) and $\mathcal{E}_N$ given by (4.3)).

Before proving this theorem we need to introduce some preliminary results.

5.1. Some preliminary results

In this section we recall de Finetti's theorem for exchangeable random variables in a setting that is useful for our aims and we discuss some related technical questions. Hereafter we use the following notation: If μ is a probability measure on $\mathbb{R}^n$ we denote by $\mu^{\otimes k}$ the probability measure on $\mathbb{R}^{nk}$ given by

$$\mu^{\otimes k} = \underbrace{\mu \otimes \cdots \otimes \mu}_{k \text{ times}}.$$

We adopt a similar notation for functions.

Definition 26. Let $\{\xi_i\}_{i\in\mathbb{N}}$ be a sequence of random variables such that each ξ_i lives in $\mathbb{R}^n$. We say that the sequence $\{\xi_i\}_{i\in\mathbb{N}}$ is exchangeable if for any finite permutation $\mathfrak{p} : \mathbb{N} \to \mathbb{N}$ we have that $\{\xi_{\mathfrak{p}(i)}\}_{i\in\mathbb{N}}$ has the same joint probability law of $\{\xi_i\}_{i\in\mathbb{N}}$.

Proposition 27 (de Finetti theorem). *Let $\{\xi_i\}_{i\in\mathbb{N}}$ be a sequence of random variables on $\mathbb{R}^n$. They are exchangeable random variables if and only if there exists a random measure ν taking values on $\mathcal{P}(\mathbb{R}^n)$ such that*

$$\mathbb{P}[(\xi_{i_1}, ..., \xi_{i_k})|\nu] = \nu^{\otimes k},$$

for any $k \in \mathbb{N}$ and $i_1, ..., i_k \in \mathbb{N}$ such that $i_j \neq i_\ell$. $\mathbb{P}[(\xi_{i_1}, ..., \xi_{i_k})|\nu]$ is by definition the conditional probability law of $(\xi_{i_1}, ..., \xi_{i_k})$ given the random measure ν.

Proof. The definitions and the proof can be found in [42, Theorem 1.1]. $\qquad\square$

Remark 28. A consequence of the de Finetti theorem is the following. If $f : \mathbb{R}^{nk} \to \mathbb{R}$ is a bounded measurable function then

$$\mathbb{E}[f(\xi_1, ..., \xi_k)] = \int_{\mathcal{P}(\mathbb{R}^n)} \int_{\mathbb{R}^{kn}} f(y_1, ..., y_k)\mu(dy_1) \cdots \mu(dy_k)\mathbb{P}_\nu(d\mu),$$

where $\mathbb{P}_\nu$ is the probability law of ν on $\mathcal{P}(\mathbb{R}^n)$.

De Finetti theorem is in general not true for *finite sequences* $\{\xi_i^N\}_{i\leq N}$ of exchangeable random variables on $\mathbb{R}^n$. On the other hand we can take advantage of a limit result as follows. First we introduce the empirical measure associated with the finite sequence $\{\xi_i^N\}_{i\leq N}$ defined as:

$$\nu_N(dx) = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_i^N}(dx).$$

Proposition 29. *Let $\{\xi_i^N\}_{i\leq N} \in \mathbb{R}^{Nn}$ be a finite sequence of exchangeable random variables on $\mathbb{R}^n$. The sequence $\{\xi_i^N\}_{i\leq N}$ converges in distribution to an infinite sequence of exchangeable random variables $\{\xi_i\}_{i\in\mathbb{N}}$ if and only if one of the following equivalent conditions hold as $N \to \infty$:*

i $(\xi_1^N, ..., \xi_k^N) \to (\xi_1, ..., \xi_k)$ in distribution and for any $k \in \mathbb{N}, k < N$,

ii $(\xi_1^N, ..., \xi_k^N, \nu_N) \to (\xi_1, ..., \xi_k, \nu)$ *in distribution and for any* $k \in \mathbb{N}, k < N$.

Proof. The proof can be found in [42, Theorem 3.2]. $\qquad\square$

Let us recall the definition of the Fisher information associated to a probability measure with density ρ_N on $\mathbb{R}^{nN}$ (see. e.g. [39]).

Definition 30. *For* $\rho_N \in W^{1,1}(\mathbb{R}^{nN})$ *we put*

$$I_N(\rho_N) := \int_{\mathbb{R}^{nN}} \frac{|\nabla\rho_N|^2}{\rho_N}$$

otherwise we set $I_N(\rho_N)$ *to be equal to* $+\infty$. *We consider the normalized Fisher information* $\mathcal{I}_N := \frac{1}{N}I_N$

Hereafter if ρ_N is a probability density on $\mathbb{R}^{Nn}$ we denote by $\rho_N^{(k)}$ the projection of ρ_N on the first k coordinates namely

$$\rho_N^{(k)}(x_1, ..., x_k) = \int_{\mathbb{R}^{N-k}} \rho_N(x_1, ..., x_k, y_{k+1}, ..., y_N) dy_{k+1} \cdots dy_N.$$

Proposition 31. *Let* ρ_N *be a probability density on* $\mathbb{R}^{Nn}$ *which is invariant with respect to coordinates permutations, then we have:*

  i I_N *(and so* $\mathcal{I}_N$*) is a proper (in the sense of having compact sublevels), convex, lower semicontinuous (l.s.c.) functional (in the sense of the weak convergence of measures on* $\mathcal{P}(\mathbb{R}^{nN})$*);*
 ii *for* $1 \le \ell \le N, \quad \mathcal{I}_\ell(\rho_N^{(\ell)}) \le \mathcal{I}_N(\rho_N)$;
iii *the (non normalized) Fisher information is super-additive, i.e., for any* $\ell = 1, ..., N$:

$$I_N(\rho_N) \ge I_\ell(\rho_N^{(\ell)}) + I_{N-l}(\rho_N^{(N-\ell)})$$

  with (in the case $I_\ell(\rho_N^{(\ell)}) + I_{N-\ell}(\rho_N^{(N-\ell)}) < +\infty$*) equality if and only if* $\rho_N = \rho_N^{(\ell)}\rho_N^{(N-\ell)}$;
 iv *if* $I(\rho_N^{(1)}) < +\infty$, *the equality* $\mathcal{I}_1(\rho_N^{(1)}) = \mathcal{I}_N(\rho_N)$ *holds if and only if* $\rho_N = (\rho_N^{(1)})^{\otimes N}$.

Proof. The proof can be found, e.g., in [39] Lemma 3.5, Lemma 3.6 and Lemma 3.7. $\qquad\square$

We conclude this section by proving some useful results about the derivative of the infimum of a family of functions and about the derivatives of convex functions.

Lemma 32. *Let* $F : \mathcal{X} \times \mathbb{I} \to \mathbb{R}$ *be a continuous function which is differentiable with respect to* $t \in \mathbb{I}$, *where* $\mathbb{I} \subset \mathbb{R}$ *is an open set and* $\mathcal{X}$ *is a metrizable compact space. Introducing* $V(t) = \min_{x \in \mathcal{X}} F(x, t)$, $t \in \mathbb{I}$, *let us suppose that* $\sup_{(x,t) \in \mathcal{X} \times \mathbb{I}} |\partial_t F(x,t)| < +\infty$, *that* $\partial_t F$ *is continuous, and that there exists a unique* $x^*(t)$ *such that* $F(x^*(t), t) = V(t)$. *Then the map* $t \to x^*(t)$ *is continuous,* V *is* $C^1(\mathbb{I})$ *and*

$$V'(t) = \partial_t F(x^*(t), t), \ t \in \mathbb{I}. \tag{5.4}$$

Proof. By Berge Maximum theorem (see, e.g., [4, Theorem 17.31]) under the hypotheses of the theorem, V is continuous and $x^*(t)$ is an upper semicontinuous correspondence. Since $x^*(t)$ is a single value correspondence (namely a function), this implies that $x^*(t)$ is continuous (see, e.g., [4, Theorem 17.6]). On the other hand, by [57, Theorem 3], we have that F is right and left differentiable and

$$V'_\pm(t_0) = \lim_{t \to t_0^\pm} \partial_t F(x^*(t), t_0).$$

Since both $\partial_t F$ (by hypothesis) and x^* (as shown above) are continuous, we have that V is differentiable and equation (5.4) holds. $\qquad\square$

Lemma 33. *Let* $F_n : \mathbb{I} \to \mathbb{R}$, *where* $\mathbb{I} \subset \mathbb{R}$ *is an open set, be a sequence of* $C^1(\mathbb{I})$ *concave functions converging point-wise as* $n \to \infty$ *to the* $C^1(\mathbb{I})$ *concave function* $F : \mathbb{I} \to \mathbb{R}$. *Then we have*

$$\lim_{n \to +\infty} F'_n(t) = F'(t), \ t \in \mathbb{I}.$$

Proof. For any $t_0 \in \mathbb{I}$ and any $\epsilon > 0$ there is $h_0 > 0$ such that for any $0 < h \le h_0$ we have

$$\frac{F(t_0) - F(t_0 - h)}{h} \le F'(t_0) + \epsilon. \tag{5.5}$$

On the other hand by the concavity of F_n we have

$$F_n'(t_0) \leq \frac{F_n(t_0) - F_n(t_0 - h)}{h}. \tag{5.6}$$

Taking the limit as $n \to +\infty$ in (5.6) and introducing the result in (5.5) we obtain $\limsup_{n \to +\infty} F_n'(t_0) \leq F'(t_0) + \epsilon$, that implies, by the arbitrary choice of ϵ,
$\limsup_{n \to +\infty} F_n'(t_0) \leq F'(t_0)$. Using a similar reasoning we are able to prove that $\liminf_{n \to +\infty} F_n'(t_0) \geq F'(t_0)$ from which we get the thesis. $\square$

5.2. Proof of Theorem 25

We start by proving three lemmas. Let us denote by $\rho_{0,N}$ the probability density which is the minimizer of the function $\mathcal{E}_N(\rho)$ and let us consider a finite sequence of random variables $(\xi_1^N, ..., \xi_N^N) \in \mathbb{R}^{Nn}$ having probability density $\rho_{0,N}$.

Lemma 34. *Under the hypotheses of Theorem 25 we have $\mathcal{E}_N(\rho_{0,N}) \leq \mathcal{E}(\rho_0)$ for any $N \in \mathbb{N}$. Furthermore the sequence $\{\xi_i^N\}_{i \leq N}$ is a sequence of exchangeable random variables such that the corresponding sequence of probability distributions is tight and converges, as $N \to +\infty$, in distribution (up to passing to a subsequence) to some infinite sequence of exchangeable random variables $\{\xi_i\}_{i \in \mathbb{N}}$.*

Proof. The first thesis of the lemma follows from the following inequalities

$$\mathcal{E}_N(\rho_{0,N}) \leq \mathcal{E}_N(\rho_0^{\otimes N}) = \mathcal{E}(\rho_0),$$

where we used the fact that $\mathcal{E}_N(\rho^{\otimes N}) = \mathcal{E}(\rho)$ for any probability density on $\mathbb{R}^n$.
First we note that by Remark 22, $\rho_{0,N}$ is unique and so it is symmetric with respect to permutations of coordinates. This means that $\{\xi_i^N\}_{i \leq N}$ are exchangeable random variables. We note that

$$\mathcal{E}_N(\rho_{0,N}) = \frac{1}{2}\mathcal{I}_N(\rho_{0,N}) + \mathcal{E}_{P,N}(\rho_{0,N}) \geq \frac{1}{2}\mathcal{I}_N(\rho_{0,N}) - C,$$

for some constant $C \geq 0$, where we used that, by Hypothesis $\mathcal{V}ii$, $\mathcal{V}(x, \mu)$ is uniformly bounded from below. Using Proposition 31 ii and the inequality $\mathcal{E}_N(\rho_{0,N}) \leq \mathcal{E}(\rho_0)$ we have

$$\mathcal{I}_k(\rho_{0,N}^{(k)}) \leq \mathcal{I}_N(\rho_{0,N}) \leq 2\mathcal{E}(\rho_0) + 2C.$$

By the fact that $\mathcal{I}_k$ is proper with respect to weak convergence of measures (see Proposition 31 i), we have that $\rho_N^{(k)}$ is a sequence of tight probability densities on $\mathbb{R}^{nk}$. Using a diagonalization argument there are a subsequence N_j and a sequence of exchangeable and compatible probability measures $\mu_\infty^{(k)}$ on $\mathcal{P}(\mathbb{R}^{nk})$ (i.e. they are such that the restriction on the first k coordinates of $\mu_\infty^{(k')}$ is exactly $\mu_\infty^{(k)}$ for any $k \leq k' \in \mathbb{N}$) such that as $j \to +\infty$

$$\rho_{0,N_j}^{(k)}(y)dy \to \mu_\infty^{(k)}, \ y \in \mathbb{R}^{nk}$$

weakly. Since $\mu_\infty^{(k)}$ are compatible and invariant with respect to coordinates permutations, by Kolmogorov's extension theorem (see, e.g., [41, Theorem 5.16]), there is a sequence of exchangeable random variables $\{\xi_i\}_{i \in \mathbb{N}}$ such that $(\xi_1, ..., \xi_k)$ has the law $\mu_\infty^{(k)}$. By Proposition 29 i, $\{\xi_i^{N_j}\}_{i \leq N}$ is defined up to N_j and converges in distribution to $\{\xi_i\}_{i \in \mathbb{N}}$, as $j \to +\infty$. $\square$

Lemma 35. *Under the hypotheses of Theorem 25, we have $\mathfrak{J}_N \to \mathfrak{J}$, as $N \to +\infty$.*

Proof. Since by Lemma 34, we have $\mathcal{E}_N(\rho_{0,N}) \leq \mathcal{E}(\rho_0)$, in order to prove that $\mathcal{E}_N(\rho_{0,N}) \to \mathcal{E}(\rho_0)$ as $N \to \infty$ it is sufficient to establish a lower bound for $\liminf_{N \to +\infty} \mathcal{E}_N(\rho_{0,N})$. Passing to a suitable subsequence we can suppose that

$$\liminf_{N \to +\infty} \mathcal{E}_N(\rho_{0,N}) = \lim_{N \to +\infty} \mathcal{E}_N(\rho_{0,N}).$$

Let ξ_i^N and ξ_i be as in Lemma 34. Then by Lemma 34, by Proposition 29 ii, by Skorohod representation theorem (see, e.g. [41, Theorem 3.2]) and using an abuse of notation identifying the subsequence with the whole sequence, we can suppose that $(\{\xi_i^N\}_{i \leq N}, \nu_N)$ converges to $(\{\xi_i\}, \nu)$ almost surely, as $N \to +\infty$. We have that

$$\mathcal{E}_N(\rho_{0,N}) = \frac{1}{2}\mathcal{I}_N(\rho_{0,N}) + \mathbb{E}[\mathcal{V}(\xi_1^N, \tilde{\nu}_N)]$$

where $\tilde{\nu}_N = \frac{1}{N-1}\sum_{2 \leq i \leq N} \delta_{\xi_i^N}$. By Proposition 31 *ii* and lower semicontinuity of Fisher information (see Proposition 31 *i*) we have that

$$\liminf_{N \to +\infty} \mathcal{I}_N(\rho_{0,N}) \geq \liminf_{N \to +\infty} \mathcal{I}_1(\rho_{0,N}^{(1)}) \geq \mathcal{I}_1(\mathrm{Law}(\xi_1)) = \mathcal{E}_K(\mathbb{E}[\nu]). \tag{5.7}$$

Since $\tilde{\nu}_N - \nu_N$ converges to 0 in total variation, by Fatou lemma, Hypothesis $\mathcal{V}i$ and Jensen inequality, we have that

$$\begin{aligned}
\liminf_{N \to +\infty} \mathbb{E}[\mathcal{V}(\xi_1^N, \tilde{\nu}_N)] &\geq \mathbb{E}[\liminf_{N \to +\infty} \mathcal{V}(\xi_1^N, \tilde{\nu}_N)] \\
&\geq \mathbb{E}[\mathcal{V}(\xi_1, \nu)] = \mathbb{E}[\mathbb{E}[\mathcal{V}(\xi_1, \nu)|\nu]] \\
&= \mathbb{E}[\tilde{\mathcal{V}}(\nu)] \geq \tilde{\mathcal{V}}(\mathbb{E}[\nu]) = \mathcal{E}_P(\mathbb{E}[\nu])
\end{aligned} \tag{5.8}$$

From the previous inequalities and the fact that ρ_0 is the minimizer, we obtain that

$$\lim_{N \to +\infty} \mathcal{E}_N(\rho_{0,N}) \geq \mathcal{E}(\mathbb{E}[\nu]) \geq \mathcal{E}(\rho_0),$$

and this concludes the proof of Lemma 35. $\qquad\square$

Let us return to the proof of Theorem 25. The first statement (5.1) is an immediate consequence of Lemma 35. In order to prove the remaining relations (5.2) and (5.3) we want to use a variational argument proposed, for example, in [26] (see also [1]). The main idea is to introduce some modified functionals $\mathfrak{E}(\lambda, \phi^2)$, $\mathfrak{E}'(\lambda', \phi^2, f)$, $\mathfrak{E}_N(\lambda, \phi^2)$ and $\mathfrak{E}'_N(\lambda', \phi^2, f)$ (where λ, λ' are real parameters and f is a suitable smooth function) such that $\mathfrak{E}(1, \phi^2) = \mathcal{E}(\phi^2)$, $\mathfrak{E}'(0, \phi^2, f) = \mathcal{E}(\phi^2)$, $\mathfrak{E}_N(1, \phi^2) = \mathcal{E}_N(\phi^2)$ and $\mathfrak{E}'_N(0, \phi^2, f) = \mathcal{E}_N(\phi^2)$, and also the derivatives with respect to λ and λ' are equal to the expressions involved in equations (5.2) and (5.3). We will prove that $\mathfrak{E}_N \to \mathfrak{E}$ and $\mathfrak{E}'_N \to \mathfrak{E}'$ for any λ, λ' and f in suitable sets, as $N \to +\infty$. This implies the convergence of the derivatives of $\mathfrak{E}_N$ and $\mathfrak{E}'_N$ with respect to λ and λ' giving us the limits (5.2) and (5.3).
More precisely, we introduce a little modification of the functionals $\mathcal{E}_N$ and $\mathcal{E}$ by writing

$$\mathfrak{E}(\lambda, \phi^2) = \int_{\mathbb{R}^n} \left(\frac{|\nabla \phi(x)|^2}{2} + \lambda \mathcal{V}(x, \phi^2)\phi^2(x) \right) dx$$

$$\mathfrak{E}'(\lambda', \phi^2, f) = \int_{\mathbb{R}^n} \left(\frac{|\nabla \phi(x)|^2}{2} + \mathcal{V}(x, \phi^2)\phi^2(x) + \lambda' f(x)V(x)\phi^2(x) \right) dx$$

$$\mathfrak{E}_N(\lambda, \phi_N^2) = \frac{1}{N} \int_{\mathbb{R}^{Nn}} \left(\frac{|\nabla \phi(x)|^2}{2} + \lambda \mathcal{V}_N(x)\phi_N^2(x) \right) dx$$

$$\mathfrak{E}'_N(\lambda', \phi_N^2, f) = \frac{1}{N} \int_{\mathbb{R}^{Nn}} \left(\frac{|\nabla \phi(x)|^2}{2} + \mathcal{V}_N(x)\phi_N^2(x) + \lambda' \sum_{i=1}^N f(x_i)V(x_i)\phi^2(x) \right) dx,$$

where $f \in C_b^\infty(\mathbb{R}^n)$ and such that $\|f\|_\infty \leq 1$. If $\lambda \in \mathbb{I} \subset \mathbb{R}$ and $\lambda' \in \mathbb{I}'$, where $\mathbb{I}$ and $\mathbb{I}'$ are small enough neighborhoods of 1 and 0 respectively, $\lambda \mathcal{V}$, $\lambda \mathcal{V}_N$, $\mathcal{V} + \lambda' fV$ and $\mathcal{V} + \lambda' \sum_i fV$ satisfy hypotheses $\mathcal{V}$ and $C\mathcal{V}$ whenever $\mathcal{V}$ and $\mathcal{V}_N$ satisfy hypotheses $\mathcal{V}$ and $C\mathcal{V}$. By Lemma 14 and Lemma 21, this means that there exist some uniquely determined positive functions $\phi_0^\lambda, \phi_0'^{\lambda'} \in H^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n, V(x)dx)$ and $\phi_{0,N}^\lambda, \phi_{0,N}'^{\lambda'} \in H^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^{nN}, \sum_{i=1}^N V(x_i)dx)$ which are the minimizers of $\mathfrak{E}(\lambda, \cdot^2)$, $\mathfrak{E}'(\lambda', \cdot^2)$, $\mathfrak{E}_N(\lambda, \cdot^2)$ and $\mathfrak{E}'_N(\lambda', \cdot^2)$ under the conditions, respectively, $\int_{\mathbb{R}^n} \phi_0^\lambda(x)^2 dx = 1$, $\int_{\mathbb{R}^n} \phi_0'^{\lambda'}(x)^2 dx = 1$, $\int_{\mathbb{R}^{nN}} \phi_{0,N}^\lambda(x)^2 dx = 1$ and $\int_{\mathbb{R}^{nN}} \phi_{0,N}'^{\lambda'}(x)^2 dx = 1$.

Lemma 36. *Under the hypotheses of Theorem 25, there are $\mathcal{X}$, $\mathcal{X}'$, $\mathcal{X}_N$ and $\mathcal{X}'_N$, that are compact subsets of $H^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n, V(x)dx)$ and $H^1(\mathbb{R}^{nN}) \cap L^2(\mathbb{R}^{nN}, \sum_{i=1}^N V(x_i)dx)$ respectively, such that $\phi_0^\lambda \in \mathcal{X}$, $\phi_0'^{\lambda'} \in \mathcal{X}'$, $\phi_{0,N}^\lambda \in \mathcal{X}_N$ and $\phi_{0,N}'^{\lambda'} \in \mathcal{X}'_N$ for any $\lambda \in \mathbb{I}$ and $\lambda' \in \mathbb{I}'$ (with ϕ_0^λ and $\phi_0'^{\lambda'}$ as defined just before the statement of this lemma).*

Proof. We give the proof only for ϕ_0^λ, the proof for $\phi_0'^{\lambda'}$, $\phi_{0,N}^\lambda$ and $\phi_{0,N}'^{\lambda'}$ being completely analogous.
By Lemma 14 we have that ϕ_0^λ satisfies the equation

$$-\Delta \phi_0^\lambda(x) + \lambda V(x)\phi_0^\lambda(x) = \mu_{0,\lambda}\phi_0^\lambda(x) - \lambda \left(2(\partial_\mu \tilde{\mathcal{V}})(x, (\phi_0^\lambda)^2) - V(x)\right)\phi_0^\lambda(x), \tag{5.9}$$

where $\mu_{0,\lambda}$ is given by expression (3.11). By Hypotheses $\mathcal{V}ii$ and $\mathcal{V}iii$ we have that $\left(2(\partial_\mu \tilde{\mathcal{V}})(x, (\phi_0^\lambda)^2) - V(x)\right)$ is bounded from below. Writing

$$E = \sup_{\lambda \in \mathbb{I}} \mathfrak{E}(\lambda, (\phi_0^\lambda)^2),$$

which is finite for $\mathbb{I}$ small enough, by multiplying equation (5.9) by $V(x)\phi_0^\lambda$ and integrating, using integration by parts and formula (3.11), we obtain

$$\int_{\mathbb{R}^n} V(x)|\nabla\phi_0^\lambda(x)|^2 dx + \int_{\mathbb{R}^n} \phi_0^\lambda(x)(\nabla V(x) \cdot \nabla\phi_0^\lambda(x)) dx + \lambda \int_{\mathbb{R}^n} (V(x)\phi_0^\lambda(x))^2 dx \leq E + C_{V,\mathcal{V}} \qquad (5.10)$$

for some constant $C_{V,\mathcal{V}}$ depending on V. Exploiting the properties (2.3) for V, a weighted Young inequality on $\phi_0^\lambda|\nabla\phi_0^\lambda|$, the fact that $-kV(x) + V(x)^2 \geq k'V(x)^2 - k''$ for any $k \in \mathbb{R}_+$ and some k', k'' depending on k and V, and multiplying both sides of (5.10) by a suitable constant we obtain that

$$\int_{\mathbb{R}^n} V(x) \left(|\nabla\phi_0^\lambda(x)|^2 + V(x)(\phi_0^\lambda(x))^2\right) dx \leq C_{\mathbb{I},V,\mathcal{V}}(E+1), \qquad (5.11)$$

where $C_{\mathbb{I},V,\mathcal{V}}$ is a positive constant depending only on $\mathbb{I}$, V and $\mathcal{V}$. By multiplying equation (5.9) by $V(x)^2\phi_0^\lambda$ and $V(x)\Delta\phi_0^\lambda$, using a similar reasoning and inequality (5.11) we obtain

$$\int_{\mathbb{R}^n} V(x)^2 \left(|\nabla\phi_0^\lambda(x)|^2 + V(x)(\phi_0^\lambda(x))^2\right) dx \leq C'_{\mathbb{I},V,\mathcal{V}}(E+1)$$

$$\int_{\mathbb{R}^n} V(x)(\Delta\phi_0^\lambda(x))^2 dx \leq C''_{\mathbb{I},V,\mathcal{V}}(E+1)^2,$$

for some positive constants $C'_{\mathbb{I},V,\mathcal{V}}$, $C''_{\mathbb{I},V,\mathcal{V}}$ depending only on $\mathbb{I}$, V and $\mathcal{V}$.

Using the fact that, by the properties (2.3) of V, $\int_{\mathbb{R}^n} V(x)((\Delta\phi(x))^2 + V(x)\phi(x)^2) dx$ is an equivalent norm of $H^2(\mathbb{R}^n, V(x)dx) \cap L^2(\mathbb{R}^n, V(x)^2 dx)$ (see, e.g. [67, Section 5.1.5] where this assertion is proven for more general Besov spaces, see also [68; 69]) we get that ϕ_0^λ is contained in some bounded subset $\mathcal{X}$ of $H^2(\mathbb{R}^n, V(x)dx) \cap L^2(\mathbb{R}^n, V(x)^2 dx)$. Since V grows to $+\infty$ when $|x| \to +\infty$, the embedding of $H^2(\mathbb{R}^n, V(x)dx) \cap L^2(\mathbb{R}^n, V(x)^2 dx)$ in $H^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n, V(x)dx)$ is compact which implies that $\mathcal{X}$ is compact in $H^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n, V(x)dx)$. $\qquad\square$

Proof of the equalities (5.2) *and* (5.3) *in Theorem 25.* We are now able to prove equation (5.2) by establishing that $\mathcal{E}_{P,N}(\rho_{0,N}) \to \mathcal{E}_P(\rho_0)$, as $N \to +\infty$. We introduce the functions

$$E_N(\lambda) = \mathfrak{E}_N(\lambda, (\phi_{0,N}^\lambda)^2) = \min_{\phi_N \in \mathcal{X}_N} \mathfrak{E}_N(\lambda, (\phi_N)^2)$$

$$E(\lambda) = \mathfrak{E}(\lambda, (\phi_0^\lambda)^2) = \min_{\phi \in \mathcal{X}} \mathfrak{E}(\lambda, \phi^2),$$

where $\mathcal{X}_N$ and $\mathcal{X}$ are the compact sets built in Lemma 36. By Lemma 35 we have that $E_N(\lambda) \to E(\lambda)$ for λ in a neighborhood $\mathbb{I}$ of 1 small enough, as $N \to +\infty$. Furthermore, since, by Lemma 36, we have that $\mathcal{X}_N$ and $\mathcal{X}$ are compact metrizable sets, we can apply Lemma 32 to E_N and E getting respectively

$$\partial_\lambda E_N(1) = \mathcal{E}_{P,N}(\rho_{0,N}) \qquad \partial_\lambda E(1) = \mathcal{E}_P(\rho_0).$$

On the other hand, since $\mathfrak{E}(\lambda, \phi)$ and $\mathfrak{E}_N(\lambda, \phi_N)$ are affine functions in λ, we have that E and E_N are concave functions, being the minimum of concave functions. This means that, by Lemma 33, $\partial_\lambda E_N(1) \to \partial_\lambda E(1)$, thus proving the equalities (5.2).

Applying a similar reasoning to

$$E'_N(\lambda, f) = \mathfrak{E}_N(\lambda, (\phi_{0,N}^\lambda)^2, f) = \min_{\phi_N \in \mathcal{X}'_N} \mathfrak{E}_N(\lambda, (\phi_N)^2, f)$$

$$E'(\lambda, f) = \mathfrak{E}(\lambda, (\phi_0^\lambda)^2) = \min_{\phi \in \mathcal{X}'} \mathfrak{E}'(\lambda, \phi^2, f),$$

we prove that

$$\int_{\mathbb{R}^n} V(x)f(x)\rho_{0,N}^{(1)}(x)dx \to \int_{\mathbb{R}^n} V(x)f(x)\rho_0(x)dx,$$

as $N \to +\infty$. Since f is any $C^\infty(\mathbb{R}^n)$ bounded function we have that $\rho_{0,N}^{(1)}$ converges to ρ_0 weakly in $L^1(\mathbb{R}^n, V(x)dx)$, as $N \to +\infty$, and this implies (5.3). This complete the proof of Theorem 25. $\qquad\square$

6. Convergence of the invariant measures

For positive $L^1(\mathbb{R}^{kn}, dx)$ functions ρ_1, ρ_2 define

$$H_k(\rho_1|\rho_2) := \begin{cases} \int_{\mathbb{R}^{kn}} \log\left(\frac{\rho_1(x)}{\rho_2(x)}\right)\rho_1(x)dx & \text{if } \mathrm{supp}(\rho_1) \subset \mathrm{supp}(\rho_2) \\ +\infty & \text{elsewhere} \end{cases},$$

and also, we define, with an abuse of notation, $H_k(\rho_1) := H_k(\rho_1|1)$.

The main aim of the present section is to prove the following theorem.

Theorem 37. *Under Hypotheses $\mathcal{V}$, CV and QV and using the notations of Theorem 25, we have*

$$\frac{1}{N}H_N(\rho_{0,N}|\rho_0^{\otimes N}) \to 0, \tag{6.1}$$

as $N \to +\infty$.

Remark 38. Thanks to Lemma 55 below, Theorem 37 implies that

$$H_k(\rho_{0,N}^{(k)}|\rho_0^{\otimes k}) \to 0$$

as $N \to +\infty$.

The proof of Theorem 37 itself will be given in Section 6.2 below. First we prove the Kac's chaoticity of the sequence $\rho_{0,N}$.

6.1. Kac's chaoticity of the sequence $\rho_{0,N}$

In this subsection we prove that the sequence of symmetric measures $\mu_{0,N}(dx_1\cdots dx_N) = \rho_{0,N}(x)dx_1\cdots dx_N$ is Kac's chaotic with limit $\mu_0(dx) = \rho_0(x)dx$ in the sense of the following definition.

Definition 39. Let $\mu_1, ..., \mu_N, ...$ be a sequence of symmetric probability measures on $\mathbb{R}^n, ..., \mathbb{R}^{nN}, ...$ respectively, denote by $\mu_K^{(k)}$, for $k \leq K \in \mathbb{N}$, the projection of μ_K on $\mathbb{R}^{nk}$, and let μ be a probability measure on $\mathbb{R}^n$. We say that $\mu_1, ..., \mu_N, ...$ is Kac's chaotic with limit μ if $\mu_N^{(k)}$ weakly converges to $\mu^{\otimes k}$ on $\mathbb{R}^{nk}$ as $N \to +\infty$.

Theorem 40. *Under Hypotheses $\mathcal{V}$ and CV we have that the sequence $\{\mu_{0,N}\}_{N \in \mathbb{N}}$ is Kac's chaotic with limit μ_0.*

Proof. Here we use the concepts and notations introduced in Section 5.
Since the measures $\mu_{0,N}^{(k)}$ are tight by Lemma 34, the statement of the theorem is equivalent to proving that any convergent subsequence of $\xi^N = (\xi_1^N, ..., \xi_N^N)$ converges to a sequence of independent identically distributed random variables $(\xi_1, ..., \xi_k, ...)$ having probability law μ_0.
By Proposition 27 and Remark 28, this is equivalent to proving that $\nu = \mu_0$ almost surely or equivalently $\mathbb{P}_\nu = \delta_{\mu_0}$, where ν is the measure related to the limit exchangeable sequence $(\xi_1, ..., \xi_k, ...)$ by Proposition 27.
By Lemma 34, the statement and the proof (in particular inequalities (5.7) and (5.8)) of Lemma 35 and Proposition 31 *iii* we have

$$\begin{aligned}\mathcal{E}(\rho_0) &\geq \mathcal{I}_k\left(\mathbb{E}[\nu^{\otimes k}]\right) + \mathcal{E}_K(\mathbb{E}[\nu]) \\ &\geq \mathcal{I}_1\left(\mathbb{E}[\nu]\right) + \mathcal{E}_K(\mathbb{E}[\nu]) = \mathcal{E}(\rho_0),\end{aligned} \tag{6.2}$$

Since, by Theorem 25, $\mathbb{E}[\nu] = \mu_0$, inequality (6.2) implies that

$$\mathcal{I}_k(\mathcal{E}[\nu^{\otimes k}]) = \mathcal{I}_1(\mathcal{E}[\nu]) = \mathcal{I}_1(\mu_0). \tag{6.3}$$

On the other hand, by Proposition 31 *iii*, from equation (6.3) we get

$$\mathbb{E}[\nu^{\otimes k}] = \mu_0^{\otimes k}. \tag{6.4}$$

Since relation (6.4) is true for any $k \in \mathbb{N}$ we must have $\nu = \mu_0$ almost surely, which implies the thesis of the theorem. $\square$

6.2. Proof of Theorem 37

In order to prove Theorem 37 we introduce some concepts and some preliminary lemmas.

Lemma 41. *Under hypotheses $\mathcal{V}$, CV and QV on $\mathcal{V}$, $\tilde{\mathcal{V}}$ and V, we have that there is $R > 0$ and some constants $a_1, a_2 \in \mathbb{R}_+$ and $a_3, a_4 \in \mathbb{R}$ such that for any $x \in \mathbb{R}$ for which $|x| \geq R$ we have*

$$\exp\left(-a_1 \int_0^{|x|} \sqrt{\bar{V}(r)}dr + a_3\right) \leq \rho_0(x) \leq \exp\left(-a_2|x|^2 + a_4\right). \tag{6.5}$$

In order to prove the previous Lemma 41, we use the following Proposition 42.

Proposition 42. *Let $B_R \subset \mathbb{R}^n$ be the closed ball of radius $R > 0$ and center $0 \in \mathbb{R}^n$. Let f, g be functions smooth in $\overline{\mathbb{R}^n \backslash B_R}$ and let K, H be measurable functions defined on $\mathbb{R}^n \backslash B_R$ such that:*

   i $\Delta|f| \leq K|f|$,
  ii $\Delta|g| \geq H|g|$,
 iii $f, g \to 0$ *as* $|x| \to +\infty$,
 iv $H(x) \geq K(x) \geq 0$ *for any* $x \notin B_R$,
  v $|f(x)| \geq |g(x)|$ *for any* $x \in \partial B_R$,

then we have $|f(x)| \geq |g(x)|$ for all $x \in \overline{\mathbb{R}^n \backslash B_R}$.

Proof. The proof can be found in [71] Theorem 8. $\qquad\square$

Proof of Lemma 41. We prove only the left part of the relation (6.5), since the proof of the right part can be done in a similar way using Hypothesis $QV\,i$ (furthermore the complete proof of the right part can also be found in [71]).

By Hypotheses $\mathcal{V}$ and Lemma 14 we have

$$\Delta\phi_0(x) \leq C_1 V(x)\phi_0(x)$$

for a suitable constant $C_1 > 0$ and for $|x| \geq R$ for some $R \geq 0$. Writing $F_{a_1', a_3'}(x) = \exp\left(-a_1' \int_0^{|x|} \sqrt{\bar{V}(r)} dr + a_3'\right)$ and choosing $a_1' > 2(e_3 + 2(n-1))$ (where e_3 is the same constant in Hypothesis QV) by Hypothesis $QV\,ii$ we have

$$\Delta F_{a_1', a_3'}(x) = a_1' F_{a_1', a_3'}(x) \frac{\left(-|x|\bar{V}'(|x|) + a_1'|x|(\bar{V}(|x|))^{\frac{3}{2}} - 2(n-1)\bar{V}(|x|)\right)}{2|x|\sqrt{\bar{V}(|x|)}} \qquad (6.6)$$

$$\geq \frac{a_1'}{4} F_{a_1', a_3'}(x) V(x). \qquad (6.7)$$

Choosing $\frac{a_1'}{4} > C_3$, since, by Lemma 14, $\phi_0 > 0$, there exists an $\bar{a}_3 > 0$ (depending on R and on a_1') such that

$$\sup_{x \in \partial B_R} F_{a_1', \bar{a}_3}(x) \leq \sup_{x \in \partial B_R} \phi_0(x).$$

By Proposition 42 we have $\phi_0(x) \geq F_{a_1', \bar{a}_3}(x)$ for $x \in \overline{\mathbb{R}^n \backslash B_R}$, which implies that $F_{2a_1', 2\bar{a}_3}(x) = F_{a_1, a_3}(x) \leq \rho_0(x)$ for $x \in \overline{\mathbb{R}^n \backslash B_R}$. $\qquad\square$

Lemma 43. *Under Hypotheses $\mathcal{V}$, $C\mathcal{V}$ and QV we have that $|H_{nN}(\rho_{0,N})|, |H_n(\rho_0)| < +\infty$ and we have the convergence*

$$\frac{1}{N} H_N(\rho_{0,N}) \to H_1(\rho_0) \qquad (6.8)$$

as $N \to +\infty$.

Proof. The fact that $|H_{nN}(\rho_{0,N})|$ is finite follows from the logarithmic Sobolev inequality for the Lebesgue measure (see, i.e., Section 4.6.1 of [5]) in fact we have

$$0 \leq \frac{2}{nN} H_N(\rho_{0,N}) \leq \log\left(\frac{2}{nNe\pi} \int_{\mathbb{R}^{nN}} \frac{|\nabla\rho_{0,N}(x)|^2}{\rho_{0,N}(x)} dx\right) = \log\left(\frac{2}{ne\pi} \mathcal{I}_N(\rho_{0,N})\right),$$

where $\mathcal{I}_N$ is the normalized Fisher information introduced in Section 5. A similar inequality holds also for $|H_1(\rho_0)|$.

The convergence (6.8) follows from the following facts

   i by Theorem 40, the sequence $\mu_{0,N}(dx) = \rho_{0,N}(x)dx$ is Kac's chaotic and converging to $\mu_0(dx) = \rho_0(x)dx$, as $N \to +\infty$;
  ii by Theorem 25, in particular equation (5.2), we get $\sup_{N \in \mathbb{N}} \mathcal{I}_N(\rho_{0,N}) < +\infty$;
 iii by Hypothesis QV, we have $\sup_{N \in \mathbb{N}} \int_{\mathbb{R}^n} \frac{\|x\|^{2+\epsilon}}{N} \rho_{0,N}(x)dx < +\infty$.

Indeed the previous properties i, ii and iii of $\mu_{N,0}$ are the hypotheses of Theorem 1.4 of [39], which implies the statement (6.8). $\qquad\square$

Proof of Theorem 37. By Hypothesis Q$V$$i$, we have that, for every $x \in \mathbb{R}^n$,

$$\int_0^{|x|} \sqrt{\bar{V}(r)} dr \le |x| \sqrt{\bar{V}(|x|)} \le \sqrt{\frac{V(x) + e_2}{e_1}} \sqrt{V(x)} \le P_1 V(x) + P_2, \tag{6.9}$$

for suitable constants $P_1, P_2 > 0$, for all $x \in \mathbb{R}^n$. Furthermore by Lemma 41, inequality (6.9) and Q$V$$i$ we get

$$|\log(\rho_0(x))| \le a_1 V(x) + a_3 + a_2 |x|^2 + a_4 \le P_3 V(x) + P_4 \tag{6.10}$$

for suitable constants $P_3, P_4 > 0$. By Lemma 43 $|H_{nN}(\rho_{0,N})|$ is bounded and using inequality (6.10), we obtain

$$\int_{\mathbb{R}^n} |\log(\rho_0(x))| \rho_{0,N}^{(1)}(x) dx \le P_3 \int_{\mathbb{R}^n} V(x) \rho_{0,N}^{(1)}(x) dx + P_4 < +\infty.$$

Thus we get that

$$\frac{1}{N} H_N(\rho_{0,N} | \rho_0^{\otimes N}) = \frac{1}{N} H_N(\rho_{0,N}) - \int_{\mathbb{R}^n} \log(\rho_0(x)) \rho_{0,N}^{(1)}(x) dx. \tag{6.11}$$

Using Lemma 43, inequality (6.10) and Theorem 25, the thesis follows. □

Remark 44. Hypothesis QV is used crucially in two points of the proof of Theorem 37. The first time it is used in Lemma 41 (and for a similar reason in inequality (6.9)) in order to be able to control the growth of $|\log(\rho_0)|$ at infinity by the function $V(x)$. The second time, Hypothesis QV is exploited in the proof of Lemma 43, in particular at the place where Theorem 1.4 of [39] is cited. Indeed, the proof of Theorem 1.4 of [39] uses in an essential way the HWI inequalities (see [74] Chapter 20) controlling the relative entropy of two measures by their Wasserstein $W_{2,\mathbb{R}^n}$ distance (see Definition 47), which is finite when the second moment exists, and their Fisher information.

7. Convergence of the probability law on the path space

In this section we prove the convergence on the path space of the N-particles system control problem (4.2), when the initial condition is the invariant measure $\rho_{0,N}$, as $N \to +\infty$, to the McKean-Vlasov optimal control problem given by (1.2) and (4.2).

Hereafter we fix a constant $T > 0$ which is the final time of the considered process. Given the spaces $\Omega_T = C([0, T]; \mathbb{R}^n)$ and $\Omega_T^N = C([0, T]; \mathbb{R}^{nN})$, we denote by $\mathbb{P}_0$ the law of the solution to the SDE (1.1) at the optimal control $\alpha = \frac{\nabla \rho}{\rho}$ and with initial condition ρ_0. Moreover, we denote by $\mathbb{P}_{0,N}$ the law of the system of N interacting diffusions (4.1) at the optimal control $A_N = \frac{\nabla \rho_{0,N}}{\rho_{0,N}}$ and with initial condition $\rho_{0,N}$. We write $\mathbb{P}_{0,N}^{(k)}$ (for $N \ge k$) for the probability measure obtained by projecting $\mathbb{P}_{0,N}$ onto Ω_T^k (the path space of the first k particles).

We introduce the notion of relative entropy.

Definition 45. If $\mathbb{P}$ and $\mathbb{Q}$ are two probability laws on the same probability space Ω, such that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$, the relative entropy (Kullback–Leibler divergence) between $\mathbb{P}$ and $\mathbb{Q}$ is defined by

$$\mathcal{H}_\Omega(\mathbb{P} | \mathbb{Q}) = \int_\Omega \log\left(\frac{d\mathbb{P}}{d\mathbb{Q}}(\omega)\right) \mathbb{P}(d\omega),$$

and it is defined to be infinity if $\mathbb{P}$ is not absolutely continuous with respect to $\mathbb{Q}$.

The following result establishes a strong form of entropy chaos for the probability laws associated with the N-particles optimal control problem.

Theorem 46. *Under hypotheses $\mathcal{V}$, CV and QV we have that for all $k \in \mathbb{N}$*

$$\lim_{N \uparrow +\infty} \mathcal{H}_{\Omega_T^k}(\mathbb{P}_{0,N}^{(k)} | \mathbb{P}_0^{\otimes k}) = 0. \tag{7.1}$$

The proof of Theorem 46 will be given in Section 7.2 below. First we would like to point out some consequences of Theorem 46.

7.1. Some consequences of Theorem 46

In this section we want to discuss some consequences of Theorem 46 in particular concerning the convergence of the probability measures $\mathbb{P}_{0,N}^{(k)}$ to $\mathbb{P}_0^{\otimes k}$ with respect to the total variation metric and the Wasserstein metric on the space of (Borel) probability measures $\mathcal{P}(\Omega)$ on Ω.

Definition 47. Let Ω be a Polish space with metric d_Ω, and let us denote by $\mathcal{P}(\Omega)$ the set of probability measures on Ω. If $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\Omega)$ the total variation distance between $\mathbb{P}$ and $\mathbb{Q}$ is the following non-negative real number

$$d_{TV,\Omega}(\mathbb{P}, \mathbb{Q}) = \sup_{A \in \mathcal{B}(\Omega)} |\mathbb{P}(A) - \mathbb{Q}(A)|,$$

where $\mathcal{B}(\Omega)$ denotes the σ-algebra of Borel subsets of Ω.

If $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\Omega)$ and $p \geq 1$ we call the p-th Wasserstein distance between $\mathbb{P}$ and $\mathbb{Q}$, $W_{p,\Omega}$, the positive number

$$W_{p,\Omega}(\mathbb{P}, \mathbb{Q}) = \left(\inf \left\{ \int_{\Omega \times \Omega} d_E(x,y)^p \pi(x,y), \text{ where } \pi \in \mathcal{P}(\Omega \times \Omega) \text{ such that } P_{1,*}\pi = \mathbb{P}, P_{2,*}\pi = \mathbb{Q} \right\} \right)^{\frac{1}{p}},$$

where $P_1 : \Omega \times \Omega \to \Omega$ and $P_2 : \Omega \times \Omega \to \Omega$ are respectively the projections on the first and second component of $\Omega \times \Omega$.

Corollary 48. *Under hypotheses $\mathcal{V}$, CV and QV we have that for all $k \in \mathbb{N}$*

$$\lim_{N \uparrow +\infty} d_{TV,\Omega_T^k}(\mathbb{P}_{0,N}^{(k)}, \mathbb{P}_0^{\otimes k}) = 0. \tag{7.2}$$

Proof. The thesis is a consequence of the well known Csiszar-Kullback inequality ([29],[43]), which is valid in arbitrary Polish spaces, yielding

$$d_{TV,\Omega}(\mathbb{P}, \mathbb{Q}) \leq \sqrt{2\mathcal{H}_\Omega(\mathbb{P}|\mathbb{Q})}. \tag{7.3}$$

$\square$

Hereafter if Ω is a separable Banach space and $\mathbb{P} \in \mathcal{P}(E)$ we write

$$M_{k,\Omega}(\mathbb{P}) = \int_\Omega ||x||_\Omega^k \mathbb{P}(dx). \tag{7.4}$$

Lemma 49. *Under Hypotheses $\mathcal{V}$, CV and QV, there is $k > 2$ (depending on the constant $\epsilon > 0$ in Hypothesis QV) such that for any $h \in \mathbb{N}$*

$$\sup \left\{ \left\{ M_{k,\Omega^h}(\mathbb{P}_{0,N}^{(h)}) | N \in \mathbb{N}, N \geq h \right\}, M_{k,\Omega^h}(\mathbb{P}_0^{\otimes h}) \right\} < +\infty$$

Proof. We provide the proof for $h = 1$ and $\mathbb{P}_{0,N}$, being the general case a straightforward generalization. Consider $\epsilon_1 > 0$ which we fix later in a suitable way. By Itô formula (applied to $|X_t^1|^2$), Jensen inequality, Doob inequality and Young inequality we have

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}_{0,N}} \left[\left(\sup_{t \in [0,T]} |X_t^1|^2 \right)^{1+\epsilon_1} \right] &\leq 2^{1+\epsilon_1} \mathbb{E}_{\mathbb{P}_{0,N}} \left[\sup_{t \in [0,T]} \left| |X_t^1|^2 - 2\int_0^t A_N^1(X_s^N) \cdot X_s^1 ds - 2t \right|^{1+\epsilon_1} \right] + \\
&\quad + 2^{2+\epsilon_1} \mathbb{E}_{\mathbb{P}_{0,N}} \left[\left(\int_0^T |A_N^1(X_s)||X_s^1| ds \right)^{1+\epsilon_1} \right] + 2^{2+\epsilon_1} T^{1+\epsilon_1} \\
&\leq 2^{2+2\epsilon_1} \mathbb{E}_{\mathbb{P}_{0,N}} [|X_0^1|^{2+2\epsilon_1}] + 2^{3+2\epsilon_1} \mathbb{E} \left[\sup_{t \in [0,T]} \left| \int_0^t X_t^1 \cdot dW_t^1 \right|^{1+\epsilon_1} \right] + \\
&\quad + 2^{2+\epsilon_1} T^{\epsilon_1} \mathbb{E}_{\mathbb{P}_{0,N}} \left[\int_0^T |A_N^1(X_s)|^{1+\epsilon_1} |X_s^1|^{1+\epsilon_1} ds \right] + 2^{2+\epsilon_1} T^{1+\epsilon_1}
\end{aligned}
$$

$$\leq \quad 2^{2+2\epsilon_1}\mathbb{E}_{\mathbb{P}_{0,N}}[|X_0^1|^{2+2\epsilon_1}] + \frac{2^{3+2\epsilon_1}(1+\epsilon_1)}{\epsilon_1}\mathbb{E}_{\mathbb{P}_{0,N}}\left[\left|\int_0^T X_t^1 \cdot dW_t^1\right|^{1+\epsilon_1}\right] +$$

$$+2^{1+\epsilon_1}(1+\epsilon_1)T^{\epsilon_1}\mathbb{E}_{\mathbb{P}_{0,N}}\left[\int_0^T |A_N^1(X_t)|^2 dt\right] + \frac{2^{3+\epsilon_1}T^{\epsilon_1}}{1-\epsilon_1} \times$$

$$\times \mathbb{E}_{\mathbb{P}_{0,N}}\left[\int_0^T |X_t^1|^{\frac{2(1+\epsilon_1)}{(1-\epsilon_1)}} dt\right] + 2^{2+\epsilon_1}T^{1+\epsilon_1}. \tag{7.5}$$

From inequality (7.5) using Jensen inequality, Itô isometry and the stationarity of the distribution of X_t we get

$$\mathbb{E}_{\mathbb{P}_{0,N}}\left[\left(\sup_{t\in[0,T]}|X_t^1|^2\right)^{1+\epsilon_1}\right] \quad \leq \quad 2^{2+2\epsilon_1}\mathbb{E}_{\mathbb{P}_{0,N}}[|X_0^1|^{2+2\epsilon_1}] + \frac{2^{3+2\epsilon_1}(1+\epsilon_1)T^{\frac{1+\epsilon_1}{2}}}{\epsilon_1}\left(\mathbb{E}_{\mathbb{P}_{0,N}}\left[|X_0^1|^2\right]\right)^{\frac{1+\epsilon_1}{2}} +$$

$$+2^{1+\epsilon_1}(1+\epsilon_1)T^{1+\epsilon_1}\mathbb{E}_{\mathbb{P}_{0,N}}\left[|A_N^1(X_0)|^2\right] + \frac{2^{3+\epsilon_1}T^{1+\epsilon_1}}{1-\epsilon_1}\mathbb{E}\left[|X_0^1|^{\frac{2(1+\epsilon_1)}{(1-\epsilon_1)}}\right] +$$

$$+2^{2+\epsilon_1}T^{1+\epsilon_1} \tag{7.6}$$

By choosing ϵ_1 small enough, i.e. such that $2\epsilon_1 < \epsilon$ and $\frac{2(1+\epsilon_1)}{1-\epsilon_1} < 2+\epsilon$ (recolling that $\epsilon > 0$ is the constant in Hypothesis QV) inequality (7.6) implies

$$\mathbb{E}_{\mathbb{P}_{0,N}}\left[\left(\sup_{t\in[0,T]}|X_t^1|^2\right)^{1+\epsilon_1}\right] \leq P_1(T,\epsilon_1)\left(\mathbb{E}_{\mathbb{P}_{0,N}}[|A_N^1(X_t)|^2 + |X_t^1|^{2+\epsilon}]\right) + P_2(T,\epsilon_1) \leq$$

$$\leq Q_1(T,\epsilon_1)\mathfrak{J}_N + Q_2(T,\epsilon_1) \tag{7.7}$$

where $P_1(T,\epsilon_1), P_2(T,\epsilon_1), Q_2(T,\epsilon_1), Q_2(T,\epsilon_1) \in \mathbb{R}_+$ are suitable functions of T and $\epsilon_1 > 0$. Since, by Theorem 25, $\sup_{N\in\mathbb{N}}\mathfrak{J}_N < +\infty$, the thesis is proved by choosing $k = 2+2\epsilon_1$. □

Corollary 50. *Under hypotheses $\mathcal{V}$, C$\mathcal{V}$ and Q$\mathcal{V}$ we have that, for all $1 \leq p \leq 2$ and $k \in \mathbb{N}$,*

$$\lim_{N\uparrow+\infty} W_{p,\Omega_T^k}(\mathbb{P}_{0,N}^{(k)}, \mathbb{P}_0^{\otimes k}) = 0. \tag{7.8}$$

Proof. The proof is based on the following inequality (see, e.g., [74] Theorem 6.15)

$$W_{p,\Omega}(\mathbb{P},\mathbb{Q}) \leq 2^{1-\frac{1}{p}}\left(\int_\Omega \|x\|_\Omega^p|\mathbb{P}-\mathbb{Q}|(dx)\right)^{\frac{1}{p}}, \tag{7.9}$$

which holds for any separable Banach space Ω and $1 \leq p$. Indeed, applying Hölder inequality and the fact that $|\mathbb{P}-\mathbb{Q}| \leq \mathbb{P}+\mathbb{Q}$ to relation (7.9), we get

$$W_{p,\Omega}(\mathbb{P},\mathbb{Q}) \leq 2^{1-\frac{1}{p}}\left(M_{k,\Omega}(\mathbb{P}) + M_{k,\Omega}(\mathbb{Q})\right)^{\frac{1}{k}}\left(d_{TV,\Omega}(\mathbb{P},\mathbb{Q})\right)^{\frac{k}{(k-p)p}}, \tag{7.10}$$

for any $p < k$. Thus, the thesis follows by applying Corollary 49 and Lemma 49 to inequality (7.10) with $\Omega = \Omega_T^k$. □

Remark 51. If we have that $V(x) \geq d_1|x|^{(2P-2)+\epsilon} - d_2$ for some constants $d_1, \epsilon > 0$, $d_2 \in \mathbb{R}$ and $P \geq 2$, then, using the techniques of the proofs of Lemma 49 and Corollary 50, it is easy to prove convergence (7.8) for any $1 \leq p \leq P$.

7.2. Proof of Theorem 46

Before giving the proof we prove some preliminary lemmas.

Lemma 52. *Under hypotheses $\mathcal{V}$, for any $N \geq 2$ and $s \geq 0$ we have*

$$\frac{1}{N}\mathbb{E}_{\mathbb{P}_{0,N}}[|A_N^1(X_s) - \alpha(X_s^1)|^2] = \int_{\mathbb{R}^{nN}}\frac{|\nabla_1\phi_{0,N}(x)|^2}{2}dx - \mu_0 +$$

$$+\int_{\mathbb{R}^{nN}}2\left(\mathcal{V}(x^1,\rho_0) - \int_{\mathbb{R}^n}\partial_\mu\mathcal{V}(y,x^1,\rho_0)\rho(y)dy\right)\phi_{0,N}^2(x)dx, \tag{7.11}$$

where μ_0 is defined in (3.11).

Proof. By a simple computation and recalling that, by Lemma 14 and Remark 15, ϕ_0 is strictly positive and C^2 we have

$$\frac{1}{N}\mathbb{E}_{\mathbb{P}_{0,N}}[|A_N^1(X_s) - \alpha(X_s^1)|^2] = \frac{1}{2}\int_{\mathbb{R}^{nN}}\left|\nabla_1\left(\frac{\phi_{0,N}}{\phi_0}\right)\right|^2\phi_0^2 dx.$$

We now prove that $\int_{\mathbb{R}^{nN}}\left|\nabla_1\left(\frac{\phi_{0,N}}{\phi_0}\right)\right|^2\phi_0^2 dx$ is finite and equal to the right hand side of equation (7.11). Let us denote by $\Psi_{R,N}$ the ground state of equation (4.5) restricted to the ball B_R, having radius R and centered in 0, with Dirichlet boundary condition (i.e. $\Psi_{R,N}$ is the solution to equation (4.5) for the minimal constant μ_N). Integrating by parts, and exploiting that $\Psi_{N,R}|_{\partial B_R} = 0$ and equation (3.10) we obtain

$$\frac{1}{2}\int_{B_R}\left|\nabla_1\left(\frac{\Psi_{N,R}}{\phi_0}\right)\right|^2\phi_0^2 dx = \int_{B_R}\left(\frac{|\nabla_1\Psi_{N,R}|^2}{2} - \frac{1}{2}\nabla_1\left(\frac{|\Psi_{N,R}|^2}{\phi_0}\right)\cdot\nabla_1\phi_0\right)dx$$

$$= \int_{B_R}\frac{|\nabla\Psi_{N,R}|^2}{2}dx - \mu_0 + \int_{B_R}2\left(\mathcal{V}(x^1,\rho_0) - \int_{\mathbb{R}^n}\partial_\mu\mathcal{V}(y,x^1,\rho_0)\rho_0(y)dy\right)|\Psi_{N,R}|^2 dx \quad (7.12)$$

Using the fact that $R\uparrow\infty$, and the density of regular functions with compact support is in $H^1(\mathbb{R}^{nN})$, we have that $\mathcal{E}_N(|\Psi_{N,R}|^2)\to\mathcal{E}_N(|\phi_{0,N}|^2)$. By exploiting a reasoning similar to the one used in the proof of Theorem 25, we prove that $|\Psi_{N,R}|^2$ converges weakly to $|\phi_{0,N}|^2$ (in $L^1(V(x)dx)$) and that $\int_{B_R}\frac{|\nabla\Psi_{N,R}|^2}{2}dx \to \int_{\mathbb{R}^{nN}}\frac{|\nabla\phi_{0,N}|^2}{2}dx$. This concludes the proof of the Lemma 52. □

Remark 53. An important consequence of Lemma 52 and Theorem 25 is that, as $N\to\infty$,

$$\frac{1}{N}\mathbb{E}_{\mathbb{P}_{0,N}}[|A_N^1(X_s) - \alpha(X_s^1)|^2] \to 0,$$

for any $s\geq 0$.

We introduce the following notation: when $\mathbb{P}$ and $\mathbb{Q}$ are defined on Ω_T^N we denote by $\overline{\mathcal{H}}_{\Omega_T^N}$ the normalized relative entropy given, for all $N\in\mathbb{N}$, by

$$\overline{\mathcal{H}}_{\Omega_T^N}(\mathbb{P}|\mathbb{Q}) := \frac{1}{N}\mathcal{H}_{\Omega_T^N}(\mathbb{P}|\mathbb{Q}). \quad (7.13)$$

The following lemma provides the expression of the normalized relative entropy in our framework.

Lemma 54. *Under the Hypotheses $\mathcal{V}$ we have that*

$$\overline{\mathcal{H}}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) = -\frac{1}{N}H_N(\rho_{0,N}|\rho_0^{\otimes N}) + \frac{1}{2}\mathbb{E}_{\mathbb{P}_{0,N}}\left[\int_0^T|A_N^1(X_s) - \alpha(X_s^1)|^2 ds\right]. \quad (7.14)$$

Proof. The proof runs in a way similar to the one performed for the Gross-Pitaevskii scaling limit in [58]. Here are the details. As a consequence of Lemma 52 we have that $\forall T > 0$

$$\mathbb{E}_{\mathbb{P}_{0,N}}\int_0^T|A_N^1(X_s)|^2 ds < +\infty \quad (7.15)$$

$$\mathbb{E}_{\mathbb{P}_{0,N}}\int_0^T|\alpha(X_s^i)|^2 ds < +\infty. \quad (7.16)$$

The inequalities (7.15) and (7.16) are *finite entropy conditions* (see, e.g. [36]) which imply that for all $T > 0$

$$\mathbb{P}_{0,N}\ll dx\otimes W^{\otimes N}, \quad \mathbb{P}_0^{\otimes N}\ll dx\otimes W^{\otimes N}$$

(where $\ll$ stands for absolute continuity and W is the law of Brownian motion on Ω_T and dx is the Lebesgue measure on the initial condition $X_0 \in \mathbb{R}^{nN}$). By applying Girsanov's theorem, we obtain in a standard way that, for all $T > 0$, the Radon-Nikodym derivative restricted to the time T is given by

$$\frac{d\mathbb{P}_{0,N}}{d\mathbb{P}_0^{\otimes N}} = \exp\left\{\log\left(\frac{\rho_0^{\otimes N}(X_0)}{\rho_{0,N}(X_0)}\right)\right.$$

$$\left. -\sum_{i=1}^N\int_0^T(A_N^i(X_s) - \alpha(X_s^i))\cdot dW_s + \frac{1}{2}\int_0^T|A_N^i(X_s) - \alpha(X_s^i)|^2 ds\right\}. \quad (7.17)$$

The relative entropy reads

$$\mathcal{H}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) =: \mathbb{E}_{\mathbb{P}_{0,N}}\left[\log\left(\frac{d\mathbb{P}_{0,N}}{d\mathbb{P}_0^{\otimes N}}\right)\right] =$$

$$= -\mathbb{E}_{\mathbb{P}_{0,N}}\left[\log\left(\frac{\rho_{0,N}(X_0)}{\rho_0^{\otimes N}(X_0)}\right)\right] + \sum_{i=1}^N \frac{1}{2}\mathbb{E}_{\mathbb{P}_{0,N}}\int_0^T |A_N^i(X_s) - \alpha(X_s^i)|^2 ds \quad (7.18)$$

Since under $\mathbb{P}_{0,N}$ the nN-dimensional process X is a solution of (4.1) with invariant probability density $\rho_{0,N}$, we get, recalling also (7.15) and (7.16), and by using the symmetry of $A_N^i(x)$ and $\rho_{0,N}$ with respect to coordinates permutations (see Remark 22)

$$\mathcal{H}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) = -\int_{\mathbb{R}^{nN}} \log\left(\frac{\rho_{0,N}(x)}{\rho_0^{\otimes N}(x)}\right)\rho_{0,N}(x)dx + \frac{1}{2}NT\int_{\mathbb{R}^{nN}} |A_N^1(x) - \alpha(x^1)|^2 \rho_{0,N}(x)dx$$

By definition of normalized relative entropy this concludes the proof of Lemma 54. $\qquad\square$

We recall an interesting property of the relative entropy in the case in which the second measure is a product measure.

Lemma 55. *We consider $\Omega = \Omega_1 \times \Omega_2$, where Ω_1 and Ω_2 are Polish spaces. Let $\mathbb{P}$ be a measure on Ω and $\mathbb{Q}_1$ and $\mathbb{Q}_2$ probability measures on Ω_1 and Ω_2 respectively. We denote by $\mathbb{Q} = \mathbb{Q}_1 \otimes \mathbb{Q}_2$ the product measure on Ω of the measures $\mathbb{Q}_1$ and $\mathbb{Q}_2$ and we suppose that $\mathbb{P} \ll \mathbb{Q}$. Then we have*

$$\mathcal{H}_\Omega(\mathbb{P}|\mathbb{Q}) \geq \mathcal{H}_{\Omega_1}(\mathbb{P}_1|\mathbb{Q}_1) + \mathcal{H}_{\Omega_2}(\mathbb{P}_2|\mathbb{Q}_2), \tag{7.19}$$

where $\mathbb{P}_1$ and $\mathbb{P}_2$ are the marginal probabilities of $\mathbb{P}$.

Proof. The proof can be found in Lemma 5.1 of [31]. $\qquad\square$

Proof of Theorem 46. We prove the statement by induction on k.

Take first $k = 1$. By applying Lemma 55 we have, for $N \geq 2$,

$$\mathcal{H}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) \geq \mathcal{H}_{\Omega_T}(\mathbb{P}_{0,N}^{(1)}|\mathbb{P}_0) + \mathcal{H}_{\Omega_T^{N-1}}(\mathbb{P}_{0,N}^{(N-1)}|\mathbb{P}_0^{\otimes(N-1)}), \tag{7.20}$$

and by repeating the same procedure we obtain

$$\mathcal{H}_{\Omega_T}(\mathbb{P}_{0,N}^{(1)}|\mathbb{P}_0) \leq \overline{\mathcal{H}}_{\Omega_T^N}(\mathbb{P}_{0,N}, \mathbb{P}_0^{\otimes N}), \tag{7.21}$$

where $\overline{\mathcal{H}}$ is the normalized entropy introduced in (7.13). Using Theorem 37, Lemma 54 and Remark 53 we have proved the thesis for $k = 1$.

For generic k, let us write $N = kN_k + r_k$, with $N_k \in \mathbb{N}, r_k = 0, ..., k-1$, and suppose that the statement is true for any $r_k < k$. By Lemma 55 we have

$$\mathcal{H}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) \geq N_k \mathcal{H}_{\Omega_T^k}(\mathbb{P}_{0,N}^{(k)}|\mathbb{P}_0^{\otimes k}) + \mathcal{H}_{\Omega_T^{r_k}}(\mathbb{P}_{0,N}^{(r_k)}|\mathbb{P}_0^{\otimes r_k}), \tag{7.22}$$

which implies:

$$\mathcal{H}_{\Omega_T^k}(\mathbb{P}_{0,N}^{(k)}|\mathbb{P}_0^{\otimes k}) \leq \frac{1}{N_k}\left\{\mathcal{H}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N}) + \mathcal{H}_{\Omega_T^{r_k}}(\mathbb{P}_{0,N}^{(r_k)}|\mathbb{P}_0^{\otimes r_k})\right\}$$

$$\leq \frac{N}{N_k}\left\{\overline{\mathcal{H}}_{\Omega_T^N}(\mathbb{P}_{0,N}|\mathbb{P}_0^{\otimes N})\right\} + \frac{1}{N_k}\mathcal{H}_{\Omega_T^{r_k}}(\mathbb{P}_{0,N}^{(r_k)}|\mathbb{P}_0^{\otimes r_k}) \quad (7.23)$$

Since when $N \uparrow \infty$ we have $\frac{N}{N_k} \to k$ and, by Theorem 37, Lemma 54 and Remark 53, $\lim_{N\uparrow+\infty} \overline{\mathcal{H}}_{\Omega_T^N}(\mathbb{P}_{0,N}, \mathbb{P}_0^{\otimes N}) \to 0$, we obtain the desired result from the induction hypothesis

$$\lim_{N\uparrow+\infty} \mathcal{H}_{\Omega_T^{r_k}}(\mathbb{P}_{N,0}^{(r_k)}|\mathbb{P}_0^{\otimes r_k}) \to 0,$$

since $r_k < k$. $\qquad\square$

8. The case of the Dirac delta potential

In this section we propose to the reader a potential $\mathcal{V}$ of the following form

$$\mathcal{V}_\delta(x, \mu) = V_0(x) + g\delta_x * \mu \tag{8.1}$$

where V_0 is a regular positive function growing at infinity, δ_x is the Dirac delta centered at $x \in \mathbb{R}^n$, $g \in \mathbb{R}_+$ is a strictly positive constant and $*$ stands for convolution. The potential $\mathcal{V}_\delta$ does not satisfies the regularity Hypothesis $\mathcal{V}i$ and $\mathcal{V}iii$. On the other hand it satisfies Hypothesis $\mathcal{V}ii$ and $C\mathcal{V}$, and (when the Gâteaux derivative is well defined) we have $\partial_\mu(\tilde{\mathcal{V}}_\delta) = 2\delta_{x-y}$, where $\tilde{\mathcal{V}}_\delta$ is defined as in (2.1), which is a positive definite distribution. A similar singular problem has been considered in the case of mean-field games in [35].

Here we do not consider the problem of proving that the optimal control ergodic problem has a unique optimal control (i.e. we do not prove here the equivalent of Theorem 17 for the potential (8.1)). We suppose here that there exists a family $\tilde{\mathcal{C}}_{V_0} \subset C^1(\mathbb{R}^n, \mathbb{R}^n)$ of vector fields α (in general we expect that it can depend on the trapping potential V_0 in (8.1), for example we can take $\tilde{\mathcal{C}}_{V_0} = \mathcal{C}_{\rho_0}$ as defined in (3.7)) such that

$$\inf_{\alpha \in \tilde{\mathcal{C}}_{V_0}} \left(\limsup_{T \to +\infty} \frac{1}{T} \left(\int_0^T \mathbb{E}_{x_0} \left[\frac{|\alpha(X_t)|^2}{2} + V_0(X_t) + g\rho_{x_0,\alpha,t}(X_t)^2 \right] dt \right) \right) =$$
$$= \mathbb{E}_{X_0 \sim \rho_0(x)dx} \left[\frac{|\nabla\rho_0(X_t)|^2}{2\rho_0^2(X_t)} + V_0(X_t) + g(\rho_0(X_t))^2 \right], \quad (8.2)$$

where $\rho_{x_0,\alpha,t}$ is the probability density of the law of the solution to the SDE (1.1) starting at $x_0 \in \mathbb{R}^n$ evaluated at time t, and ρ_0 is the density of the probability distribution minimizing the functional

$$\mathcal{E}_\delta(\rho) = \mathcal{E}_K(\rho) + \mathcal{E}_{\delta,P}(\rho) = \int_{\mathbb{R}^n} \left(\frac{|\nabla\rho(x)|^2}{\rho(x)} + V_0(x)\rho(x) + g(\rho(x))^2 \right) dx. \tag{8.3}$$

In other words we suppose that in the set $\mathcal{C}$ (introduced in Section 3.2) the optimal control for the problem (1.1) with cost functional (1.2) and potential $\mathcal{V}_\delta$ (see [3] for an alternative derivation of a stochastic process associated with the above cost functional) exists and it is given by $\alpha = \frac{\nabla\rho_0}{\rho_0}$. What we want to consider here is an N-particle problem converging to the solution of the optimal control ergodic problem just described (namely we are looking for an analogous of Theorem 46).

Obviously, since $\mathcal{V}_\delta$ is not well defined for measures μ that are not absolutely continuous measures, we consider here an approximating potential of the form

$$\mathcal{V}_{\delta,N}(x, \mu) = V_0(x) + \int_{\mathbb{R}^n} v_N(x - y)\mu(dy),$$

where $v_N : \mathbb{R}^n \to \mathbb{R}$ is a sequence of positive functions converging in the sense of distributions to a Dirac delta δ_0 when $N \to \infty$. Let us choose a specific sequence of the following form

$$v_N(x) = N^{n\beta} v_0 \left(N^\beta x \right), \ x \in \mathbb{R}^n \tag{8.4}$$

for $\beta > 0$, where v_0 is a positive smooth radially symmetric function with compact support. We take the N-particles approximation having the control $A(x_1, ..., x_N)$ given by the logarithm derivative of $\rho_{0,N}$ that is the minimal probability density of the energy functional $\mathcal{E}_\delta$ associated with $\mathcal{V}_{\delta,N}$, namely

$$\mathcal{E}_{\delta,N}(\rho) = \mathcal{E}_{K,N}(\rho) + \mathcal{E}_{\delta,P,N}(\rho) =$$
$$= \frac{1}{N} \sum_{i=1}^N \left(\int_{\mathbb{R}^{Nn}} \left(\frac{|\nabla_i\rho|^2}{\rho} + V_0(x_i)\rho \right) dx + \frac{1}{N-1} \sum_{j=1,...,N, j\neq i} \int_{\mathbb{R}^{Nn}} v_N(x_i - x_j)\rho dx \right).$$

In the rest of the paper we show how the results on Bose-Einstein condensation (mainly for $n = 3$, see, e.g., [49; 50; 52; 53; 54; 59; 65]) can be used to study the convergence of the N-particles approximation of the control problem with potential (8.1). For this reason hereafter we shall limit our discussion to the case $n = 3$.

8.1. Intermediate scaling limit

The case $0 < \beta < 1$, where β is the parameter used in the rescaling (8.4), which is known as intermediate scaling limit, is very similar to the regular case that we treated in the first part of the paper. Indeed in this case we can prove the following theorem.

Theorem 56. *Under the previous hypotheses and notations, if $0 < \beta < 1$ we have the $N \to +\infty$ convergence statements $\mathcal{E}_{\delta,N}(\rho_{0,N}) \to \mathcal{E}_\delta(\rho_0)$, $\mathcal{E}_{\delta,P,N}(\rho_{0,N}) \to \mathcal{E}_{\delta,P}(\rho_0)$ and $\rho_{0,N}^{(1)} \to \rho_0$ (where the later convergence is in the weak L^1 sense) with the constant $g = \int_{\mathbb{R}^3} v_0(x)dx$ (where $g \in \mathbb{R}_+$ is the constant appearing in equation (8.1) and (8.2)).*

Proof. The proof of the theorem can be found in [50] for $0 \le \beta < \frac{1}{3}$ (for any n and a more general class of potentials v_0 than the one considered here) and in [1] for $0 \le \beta < 1$ (for $n = 3$ and a positive-definite interaction potential v_0). □

Theorem 56 is the analogue of Theorem 25 in this context and it proves that $\mathcal{E}_{\delta,N}$ and $\mathcal{E}_\delta$ satisfy the thesis of Theorem 25. Thanks to Theorem 56 we can repeat the reasoning performed in Section 7, obtaining:

Theorem 57. *Under the previous hypotheses and notations, if $0 < \beta < 1$ we have that the law $\mathbb{P}_{0,N}^{(k)}$ of the first k particles satisfying the system (4.1), with $\mathcal{V}$ replaced by $\mathcal{V}_\delta$, converges in total variation on the path space $C([0,T], \mathbb{R}^{3k})$ to $\mathbb{P}_0^{\otimes k}$ (where $\mathbb{P}_0$ is the law on $C([0,T], \mathbb{R}^3)$ of the system (1.1) associated with (8.1)).*

Proof. The proof can be found in [1]. □

8.2. Gross-Pitaevskii scaling limit

The case $\beta = 1$ is completely different with respect to the previous ones. The main difference between the cases $0 < \beta < 1$ and $\beta = 1$ is that in this latter case the value function convergence result (5.2) does not hold.

Theorem 58. *Under the previous hypotheses and notations, if $\beta = 1$ we have that $\mathcal{E}_{\delta,N}(\rho_{0,N}) \to \mathcal{E}_\delta(\rho_0)$ and $\rho_{0,N}^{(1)} \to \rho_0$ (where the latter convergence is in the weak sense in L^1) for $g = 4\pi a$ (where $g \in \mathbb{R}_+$ is the constant appearing in equation (8.1) and (8.2), and $a > 0$ is the scattering length of the interaction potential v_0 (a sort of effective range of the interaction potential, for details see [53])). Furthermore putting $\hat{s} = \frac{1}{g} \int_{\mathbb{R}^3} \frac{|\nabla \rho_0|^2}{\rho_0} dx \in (0,1)$ we have, as $N \to +\infty$:*

$$\mathcal{E}_{K,N}(\rho_{0,N}) \to \mathcal{E}_{\delta,K}(\rho_0) + g\hat{s} \int_{\mathbb{R}^3} \rho_0^2(x)dx.$$

Proof. The proof of the first part of the theorem is a, by this time, well-known relevant result proven in [52; 54; 59]. The second part is proven in [54]. □

In this case we cannot repeat the reasoning of Section 7 since we are not able to prove that the relative entropy $\mathcal{H}(\mathbb{P}_{0,N}^{(k)}|\mathbb{P}_0^{\otimes k})$ converges to 0 (in fact we do not know whether the entropy converges to 0 or to another value). On the other hand it is possible to prove a weaker result for $\beta = 1$ (see [58] for a different kind of convergence and [72] for a transition to chaos result).

Theorem 59. *Under the previous hypotheses and notations, if $\beta = 1$ we have that the law $\mathbb{P}_{0,N}^{(k)}$ converges weakly on the path space $C([0,T], \mathbb{R}^{3k})$ to $\mathbb{P}_0^{\otimes k}$.*

Proof. The proof can be found in [2]. □

Acknowledgments

References

[1] Sergio Albeverio, Francesco C. De Vecchi, Andrea Romano, and Stefania Ugolini. Strong Kac's chaos in the mean-field Bose-Einstein condensation. *Stoch. Dyn.*, 20(5):2050031, 21, 2020.

[2] Sergio Albeverio, Francesco C. De Vecchi, and Stefania Ugolini. Entropy chaos and Bose-Einstein condensation. *J. Stat. Phys.*, 168(3):483–507, 2017.

[3] Sergio Albeverio and Stefania Ugolini. A Doob h-transform of the Gross-Pitaevskii Hamiltonian. *J. Stat. Phys.*, 161(2):486–508, 2015.

[4] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis.* Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[5] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses].* Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[6] Ari Arapostathis, Anup Biswas, and Johnson Carroll. On solutions of mean field games with ergodic cost. *J. Math. Pures Appl. (9)*, 107(2):205–251, 2017.

[7] Ari Arapostathis, Vivek S. Borkar, and Mrinal K. Ghosh. *Ergodic control of diffusion processes*, volume 143 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2012.

[8] Julio Backhoff, Giovanni Conforti, Ivan Gentil, and Christian Léonard. The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities. *Probab. Theory Related Fields*, 178(1-2):475–530, 2020.

[9] Martino Bardi and Ermal Feleqi. Nonlinear elliptic systems and mean-field games. *NoDEA Nonlinear Differential Equations Appl.*, 23(4):Art. 44, 32, 2016.

[10] Martino Bardi and Fabio S. Priuli. Linear-quadratic N-person and mean-field games with ergodic cost. *SIAM J. Control Optim.*, 52(5):3022–3052, 2014.

[11] Matteo Basei and Huyên Pham. A weak martingale approach to linear-quadratic McKean-Vlasov stochastic control problems. *J. Optim. Theory Appl.*, 181(2):347–382, 2019.

[12] Erhan Bayraktar, Andrea Cosso, and Huyên Pham. Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics. *Trans. Amer. Math. Soc.*, 370(3):2115–2160, 2018.

[13] Alain Bensoussan, Jens Frehse, and Sheung Chi Phillip Yam. The master equation in mean field theory. *J. Math. Pures Appl. (9)*, 103(6):1441–1474, 2015.

[14] Vladimir I. Bogachev, Nicolai V. Krylov, Michael Röckner, and Stanislav V. Shaposhnikov. *Fokker-Planck-Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2015.

[15] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013.

[16] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games*, volume 201 of *Annals of Mathematics Studies.* Princeton University Press, Princeton, NJ, 2019.

[17] Pierre Cardaliaguet, P. Jameson Graber, Alessio Porretta, and Daniela Tonon. Second order mean field games with degenerate diffusion and local coupling. *NoDEA Nonlinear Differential Equations Appl.*, 22(5):1287–1317, 2015.

[18] Pierre Cardaliaguet, Jean-Michel Lasry, Pierre-Louis Lions, and Alessio Porretta. Long time average of mean field games. *Netw. Heterog. Media*, 7(2):279–301, 2012.

[19] Pierre Cardaliaguet and Catherine Rainer. An example of multiple mean field limits in ergodic differential games. *NoDEA Nonlinear Differential Equations Appl.*, 27(3):Paper No. 25, 19, 2020.

[20] Eric A. Carlen. Conservative diffusions. *Comm. Math. Phys.*, 94(3):293–315, 1984.

[21] Eric A. Carlen. Existence and sample path properties of the diffusions in Nelson's stochastic mechanics. In *Stochastic processes—mathematics and physics (Bielefeld, 1984)*, volume 1158 of *Lecture Notes in Math.*, pages 25–51. Springer, Berlin, 1986.

[22] René Carmona and François Delarue. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.*, 43(5):2647–2700, 2015.

[23] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling.* Springer, Cham, 2018. Mean field FBSDEs, control, and games.

[24] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling.* Springer, Cham, 2018. Mean field games with common noise and master equations.

[25] Yongxin Chen, Tryphon T. Georgiou, and Michele Pavon. On the relation between optimal transport and Schrödinger bridges: a stochastic control viewpoint. *J. Optim. Theory Appl.*, 169(2):671–691, 2016.

[26] A. Yu Cherny and A.A. Shanenko. The kinetic and interaction energies of a trapped Bose gas: Beyond the

mean field. *Physics Letters A*, 293(5-6):287–292, 2002.

[27] Marco Cirant. A generalization of the Hopf-Cole transformation for stationary mean-field games systems. *C. R. Math. Acad. Sci. Paris*, 353(9):807–811, 2015.

[28] Ana Bela Cruzeiro, Carlos Oliveira, and Jean-Claude Zambrini. Time-symmetric optimal stochastic control problems in space-time domains. *arXiv preprint arXiv:2007.02110*, 2020.

[29] I. Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.

[30] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[31] Francesco De Vecchi and Stefania Ugolini. An entropy approach to Bose-Einstein condensation. *Commun. Stoch. Anal.*, 8(4):517–529, 2014.

[32] Fabrice Mao Djete, Dylan Possamaï, and Xiaolu Tan. Mckean-vlasov optimal control: limit theory and equivalence between different formulations. *arXiv preprint arXiv:2001.00925*, 2020.

[33] Mao Fabrice Djete. Extended mean field control problem: a propagation of chaos result. *arXiv preprint arXiv:2006.12996*, 2020.

[34] Ermal Feleqi. The derivation of ergodic mean field game equations for several populations of players. *Dyn. Games Appl.*, 3(4):523–536, 2013.

[35] Franco Flandoli, Maddalena Ghio, and Giulia Livieri. *N*-player games and mean field games of moderate interactions. *arXiv preprint arXiv:2104.03944*, 2021.

[36] Hans Föllmer. Random fields and diffusion processes. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 101–203. Springer, Berlin, 1988.

[37] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[38] Francesco Guerra and Laura M. Morato. Quantization of dynamical systems and stochastic control theory. *Phys. Rev. D (3)*, 27(8):1774–1786, 1983.

[39] Maxime Hauray and Stéphane Mischler. On Kac's chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014.

[40] Max-Olivier Hongler. Mean-field games and swarms dynamics in Gaussian and non-Gaussian environments. *Journal of Dynamics & Games*, 7(1):1, 2020.

[41] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[42] Olav Kallenberg. *Probabilistic symmetries and invariance principles*. Probability and its Applications (New York). Springer, New York, 2005.

[43] Solomon Kullback. A lower bound for discrimination information in terms of variation (corresp.). *IEEE transactions on Information Theory*, 13(1):126–127, 1967.

[44] Daniel Lacker. Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55(3):1641–1672, 2017.

[45] Daniel Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electron. Commun. Probab.*, 23:Paper No. 45, 11, 2018.

[46] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[47] Rémi Lassalle and Jean Claude Zambrini. A weak approach to the stochastic deformation of classical mechanics. *J. Geom. Mech.*, 8(2):221–233, 2016.

[48] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[49] Mathieu Lewin. Mean-field limit of Bose systems: rigorous results. *arXiv preprint arXiv:1510.04407*, 2015.

[50] Mathieu Lewin, Phan Nam, and Nicolas Rougerie. The mean-field approximation and the non-linear Schrödinger functional for trapped Bose gases. *Transactions of the American Mathematical Society*, 368(9):6131–6157, 2016.

[51] Elliott H. Lieb and Michael Loss. *Analysis*. American Mathematical Society, 2001.

[52] Elliott H. Lieb and Robert Seiringer. Proof of Bose-Einstein condensation for dilute trapped gases. *Physical review letters*, 88(17):170409, 2002.

[53] Elliott H. Lieb, Robert Seiringer, Jan Philip Solovej, and Jakob Yngvason. *The mathematics of the Bose gas and its condensation*, volume 34 of *Oberwolfach Seminars*. Birkhäuser Verlag, Basel, 2005.

[54] Elliott H. Lieb, Robert Seiringer, and Jakob Yngvason. Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional. In *The Stability of Matter: From Atoms to Stars*, pages 685–697. Springer, 2001.

[55] Luca Lorenzi and Marcello Bertoldi. *Analytical methods for Markov semigroups*, volume 283 of *Pure and*

Applied Mathematics (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2007.

[56] Toshio Mikami and Michèle Thieullen. Duality theorem for the stochastic optimal control problem. *Stochastic Process. Appl.*, 116(12):1815–1835, 2006.

[57] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

[58] Laura M. Morato and Stefania Ugolini. Stochastic description of a Bose-Einstein condensate. *Ann. Henri Poincaré*, 12(8):1601–1612, 2011.

[59] Phan Thanh Nam, Nicolas Rougerie, and Robert Seiringer. Ground states of large bosonic systems: the Gross–Pitaevskii limit revisited. *Analysis & PDE*, 9(2):459–485, 2016.

[60] Edward Nelson. *Quantum fluctuations*. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1985.

[61] Huyên Pham and Xiaoli Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55(2):1069–1101, 2017.

[62] Huyên Pham and Xiaoli Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: Control, Optimisation and Calculus of Variations*, 24(1):437–461, 2018.

[63] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[64] Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[65] Nicolas Rougerie. De Finetti theorems, mean-field limits and Bose-Einstein condensation. *arXiv preprint arXiv:1506.05263*, 2015.

[66] Nicolas Rougerie. Scaling limits of bosonic ground states, from many-body to non-linear schrödinger. *EMS Surveys in Mathematical Sciences*, 2021.

[67] Hans-Jürgen Schmeisser and Hans Triebel. *Topics in Fourier analysis and function spaces*. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1987.

[68] Thomas Schott. Function spaces with exponential weights. I. *Math. Nachr.*, 189:221–242, 1998.

[69] Thomas Schott. Function spaces with exponential weights. II. *Math. Nachr.*, 196:231–250, 1998.

[70] Laurent Schwartz. *Théorie des distributions*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Hermann, Paris, 1966. Nouvelle édition, entiérement corrigée, refondue et augmentée.

[71] Barry Simon. Pointwise bounds on eigenfunctions and wave packets in N-body quantum systems. III. *Trans. Amer. Math. Soc.*, 208:317–329, 1975.

[72] Stefania Ugolini. Bose-Einstein condensation: a transition to chaos result. *Commun. Stoch. Anal.*, 6(4):565–587, 2012.

[73] Denis Ullmo, Igor Swiecicki, and Thierry Gobron. Quadratic mean field games. *Phys. Rep.*, 799:1–35, 2019.

[74] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[75] Kunio Yasue. Quantum mechanics and stochastic control theory. *J. Math. Phys.*, 22(5):1010–1020, 1981.

[76] Kôsaku Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, sixth edition, 1980.

[77] J.-C. Zambrini. Variational processes and stochastic versions of mechanics. *J. Math. Phys.*, 27(9):2307–2330, 1986.