GENERALIZED MATHER PROBLEM AND HOMOGENIZATION OF HAMILTON-JACOBI EQUATIONS

DIOGO A. GOMES & ENRICO VALDINOCI

Abstract. We present a general approach to the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting through infinite dimensional linear programming and generalized Aubry-Mather theory.

Contents

1. Introduction

The objective of this paper is to present a general approach to the homogenization of random stationary ergodic Hamilton-Jacobi equations arising in deterministic and stochastic optimal control. In our approach we will use infinite dimensional linear programming and the generalized Aubry-Mather theory.

Our setting is the following. Let $\mathrm{Mat}(n\times n)$ be the space of $(n\times n)$ real matrices and Ω be a probability space. We consider a Hamiltonian

$$H(M,p,x,X,\epsilon,\omega):\mathrm{Mat}(n\times n)\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^+\times\Omega\to\mathbb{R}.$$

We will study the homogenization properties of the Hamilton-Jacobi equation

$$\begin{cases} -\partial_t u^\epsilon + H\left(D_x^2 u^\epsilon, D_x u^\epsilon, x, \dfrac{x}{\epsilon}, \epsilon, \omega\right) = 0 \\ u^\epsilon(x,T,\omega) \,=\, \psi(x)\,, \end{cases} \tag{1}$$

as $\epsilon \to 0$. Here above, $u^\epsilon(x,t,\omega) : \mathbb{R}^n \times [0,T] \times \Omega \longrightarrow \mathbb{R}$ and ϵ is a small, positive parameter and $\psi(x,X,\omega) : \mathbb{R}^n \times \mathbb{R}^n \times \Omega \to \mathbb{R}$ is assumed to be a bounded smooth function of the variables (x,X), and to satisfy further additional stationarity hypotheses in ω. We will prove that, under suitable assumptions, there exists a non-random function $\bar{u}(x,t)$ such that $u^\epsilon \to \bar{u}$. Furthermore, there is a function $\overline{H}(M,p,x)$, the so-called Effective Hamiltonian, non-random, such that $\bar{u}$ is a viscosity solution of

$$\begin{cases} -\partial_t \bar{u} + \overline{H}\left(D_x^2 \bar{u}, D_x \bar{u}, x\right) = 0 \\ \bar{u}(x,T) \,=\, \psi(x)\,. \end{cases} \tag{2}$$

Our strategy to study this homogenization problem is the following: first we establish some general properties of the original problem by converting it into a linear programming problem. Then, we introduce a homogenized problem. Finally, we prove the convergence of the sequence u^ϵ to a non-random limit and show that this limit agrees with the homogenized problem.
To accomplish these goals, we will take a series of assumptions that will be put forward in section 3.
Homogenization problems have been quite popular in recent years, and many steps have been made towards the understanding of these problems (see, e.g., [BLP76, LPV86, Eva92, Con96, Sou99, Alv99, Ish99, Ish00, RT00, CDI01, EG01, AB02, LS03, Gom03, LS05, KRV06, Kos06]). Several author, see [KS, CS02, MS98, GR06] have used linear programming techniques to study deterministic and stochastic optimal control problems.
We should point out that there were some known connections between homogenization theory and Aubry-Mather theory (see, for instance, [LS03]), and that convex analysis and duality methods were used in [KRV06] to prove a homogenization result related to ours. However, to our knowledge, this is first paper in which this relation is worked out systematically, specially in what concerns second-order equations.

2. The main result

We now state our main result. The assumptions under which it holds are very general and, for the convenience of the reader, their detailed statement will be postponed to section 3.

Theorem 1. *Let Assumptions 1–15 of section 3 hold and let u^ϵ be a viscosity solution of* (1). *Then:*

- *The solution u^ϵ may be represented as*

$$\begin{aligned} u^\epsilon(x,t,\omega) &:= \inf \int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} L(x',X,v,\omega)\, d\mu^\epsilon(x',X,v,s;\omega) \\ &\quad + \int_{\mathbb{R}^n\times\mathbb{R}^n} \psi(x')\, d\nu^\epsilon_T(x',X;\omega)\,, \end{aligned}$$

where the infimum is taken over all measures $(\mu^\epsilon, \nu^\epsilon_T)$ which, for every $\omega \in \Omega$, satisfy: μ^ϵ is a measure on $\mathbb{R}^n \times \mathbb{R}^n \times U \times [t,T]$ with total mass $T-t$, ν^ϵ_T is a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$, and μ^ϵ and ν^ϵ_T are such that

$$\begin{aligned} &\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} \partial_t\phi(x',X,s) + \tilde{A}\phi(x',X,s)\, d\mu^\epsilon(x',X,v,s;\omega) \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x',X,T\right)\, d\nu^\epsilon_T(x',X;\omega) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x',X,t\right)\, d\nu^\epsilon_t(x',X)\,, \end{aligned}$$

for any $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n \times [t,T])$.
- *The norms $\|u^\epsilon_t\|_\infty$ and $\|Du^\epsilon\|_\infty$ are bounded uniformly in ϵ.*
- *Up to subsequences, $u^\epsilon(x,t,\omega)$ converges to a suitable $\bar{u}(x,t)$ uniformly in $K \times [0,T] \times \Omega$, for any compact set $K \subset \mathbb{R}^n$. The function $\bar{u}$ is a viscosity solution of* (2)*, for a suitable $\overline{H}$.*

During the proof of Theorem 1 several results of independent interest will show up. For instance: a "dual representation" of u^ϵ will be provided (see Proposition 6), a way to reconstruct u^ϵ from subsolutions is found (see (52) below), convergence issues of minimizing measures are discussed (see Proposition 11), a measure theoretic construction of $\bar{u}$ is obtained (see (72) or (73), according to the proper scaling), and an Effective Lagrangian is constructed (see (59) below). Moreover, under an additional (still quite general) hypothesis (namely, Assumption 16 on page 43), the Effective Hamiltonian may be reconstructed from the Effective Lagrangian by duality (see Proposition 33).
We now list the hypotheses needed for Theorem 1 to hold, that is Assumptions 1–15, collected in the next section (the additional hypothesis in Assumption 16 on page 43 is not needed for Theorem 1, but only for Proposition 33).

3. List of assumptions

We now list the assumptions needed for our purposes. Some of these assumptions will be slightly redundant, but we prefer to list similar assumptions in a separate way any time they are needed in different parts of the paper, since some of the estimates that will be obtained may be of general interest.
Our first assumption concerns the structure of the problem, and will be assumed without explicit mention throughout the paper.

Assumption 1. We suppose that the Hamilton-Jacobi equation (1) originates from a stochastic control problem. The dynamics of this control problem is

$$d\mathbf{x} = f\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\right) dt + \sigma\left(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \epsilon, \mathbf{v}, \omega\right) dW_t\,, \tag{3}$$

where $\mathbf{v}$ is a bounded progressively measurable control, lying in a closed convex set $U \subseteq \mathbb{R}^m$, which, without loss of generality, will be assumed to contain the origin 0,

$$f : \mathbb{R}^n \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R}^n$$

is the drift coefficient,

$$\sigma : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^+ \times U \times \Omega \to \mathrm{Mat}(n \times p)$$

is the diffusion coefficient, and W_t is an p-dimensional Brownian motion defined in a probability space $\tilde{\Omega}$ which is independent of Ω. The probability on $\tilde{\Omega}$ will be denoted by P, and E will denote the corresponding expected value. We assume enough regularity and growth conditions in f and σ such that the stochastic differential equation (3) is well defined for all bounded progressively measurable controls $\mathbf{v}$.
The running cost of this control problem is given by the Lagrangian

$$L(x, X, v, \omega) : \mathbb{R}^n \times \mathbb{R}^n \times U \times \Omega \to \mathbb{R},$$

where $U \subseteq \mathbb{R}^m$ is the control space, which is taken to be closed and convex. We suppose also that L is continuous in (x, X, v) and bounded from below. Without loss of generality, we will assume that $\inf L = 0$. Furthermore, if U is unbounded we suppose that

$$\lim_{R\to+\infty} \inf_{\substack{(x,X,\omega)\\ |v|\geq R}} \frac{L(x, X, v, \omega)}{|v|} = +\infty\,. \tag{4}$$

For simplicity, we assume that the terminal cost ψ is a bounded C^∞ function with bounded derivatives (less regular data may be treated as well).
We suppose that, for $T > 0$, equation (1) possesses, for any fixed $\omega \in \Omega$, one and only one viscosity solution u^ϵ defined for all $t \in [0, T]$, which is the value function of the stochastic control problem, that is, it can be represented in the form

$$\begin{aligned} u^\epsilon(x, t, \omega) &= \\ &\inf E\left[\int_t^T L\left(\mathbf{x}(s), \frac{\mathbf{x}(s)}{\epsilon}, \mathbf{v}(s), \omega\right) ds + \psi\left(\mathbf{x}(T)\right)\right]\,, \end{aligned} \tag{5}$$

where the infimum is taken over all trajectories of the controlled dynamics (3) with initial condition $\mathbf{x}(t) = x$. ◀

We refer to [FR75, Kry80, FS06, Eva06a] for further discussions about the stochastic version of Hamilton-Jacobi equations.
The Hamiltonian

$$H(M, p, z, x, X, \epsilon, \omega) : \mathrm{Mat}(n \times n) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \Omega \to \mathbb{R}$$

is related with the Lagrangian by duality using the generalized Legendre transform

$$\begin{aligned} &H(M, p, x, X, \epsilon, \omega) \\ &= \sup_{v\in U} -f(x, X, v, \omega) \cdot p - \frac{\sigma\sigma^T(x, X, \epsilon, v, \omega)}{2} : M - L(x, X, v, \omega)\,. \end{aligned} \tag{6}$$

Here and in the sequel, we use the following notation: given $A, B \in \mathrm{Mat}(n \times n)$, we set

$$A : B = \operatorname{tr} A^T B,$$

if A and B are symmetric, and we will use the convention that for general matrices $A : B = \frac{1}{4} \operatorname{tr} (A^T + A)(B + B^T)$. Note that (6) implies that H is decreasing in M, that is, if $N \in \mathrm{Mat}(n \times n)$ is symmetric and non-negatively definite, then

$$H(M + N, p, x, X, \epsilon, \omega) \leq H(M, p, x, X, \epsilon, \omega)\,. \tag{7}$$

Assumption 2. We suppose that

$$\sup_{y} \sup_{\epsilon > 0} \sup_{x,\omega} \left| H\Big(D^2\Psi(x), D\Psi(x), x - y, \frac{x}{\epsilon}, \epsilon, \omega\Big) \right| < \infty\,, \tag{8}$$

for any $\Psi \in C^2(\mathbb{R}^n)$. ◀

We now suppose that the probability space Ω has nice slicing properties with respect to products with real spaces:

Assumption 3. If $\mu(r, \omega, r')$ is a finite measure on $\mathbb{R}^d \times \Omega \times \mathbb{R}^{d'}$, there exist a finite measure $\Theta(r')$ on $\mathbb{R}^{d'}$ and, for Θ-almost any $r' \in \mathbb{R}^{d'}$, a finite measure $\zeta(r, \omega; r')$ on $\mathbb{R}^d \times \Omega$ in such a way that

$$\mu(r, \omega, r') = \zeta(r, \omega; r')\Theta(r')\,.$$

Analogously, if $\mu(r, \omega)$ is a finite measure on $\mathbb{R}^d \times \Omega$, there exist a finite measure $\Theta(\omega)$ on Ω and, for Θ-almost any $\omega \in \Omega$, a finite measure $\zeta(r; \omega)$ on $\mathbb{R}^d$ in such a way that

$$\mu(r, \omega) = \zeta(r; \omega)\Theta(\omega)\,.$$

◀

Assumption 3 is fulfilled, for instance, if $\Omega = \mathbb{R}^m$: see Theorem 10 on page 14 of [Eva90]. The next two assumptions will be used in section 4 and will be assumed from that point on.

Assumption 4. Suppose that the Lagrangian L, the drift f and the diffusion coefficient σ satisfy

$$|f(x', X, v, \omega)|^p + \frac{|\sigma\sigma^T(x', X, \epsilon, v, \omega)|^p}{(1 + |x'|^2)^{p/2}} \leq C\Big(1 + L(x', X, v, \omega)\Big)$$

for any $(x', X, v, \omega) \in \mathbb{R}^n \times \mathbb{R}^n \times U \times \Omega$ and any $\epsilon \in \mathbb{R} \setminus \{0\}$, for suitable $C > 0$ and $p > 1$. ◀

Assumption 5. We assume that $L(x, X, 0, \omega)$ is uniformly bounded. ◀

As a notational remark, we point out that, in most of this paper, we think x to be the "frozen" initial datum of the controlled dynamics (and so the "frozen" variable of u^ϵ), while x' will denote the variable of the integration of the measure we will construct.
The next assumption is quite natural and requires that the viscosity solutions can be approximated by smooth subsolutions. This assumption will be required from section 6 onwards.

Assumption 6. For any $\delta > 0$ there exists $v_\delta^\epsilon \in C^2(\mathbb{R}^n \times [t,T] \times \Omega)$ satisfying

$$\partial_t v_\delta^\epsilon(x,t,\omega) - H\left(D_x^2 v_\delta^\epsilon(x,t,\omega), D_x v_\delta^\epsilon(x,t,\omega), \frac{x}{\epsilon}, \omega\right) + \delta \geq 0\,, \tag{9}$$

$v_\delta^\epsilon(x,T,\omega) = \psi(x)$, and

$$\lim_{\delta \to 0^+} v_\delta^\epsilon = u^\epsilon \tag{10}$$

uniformly in $(x,t,\omega) \in \mathbb{R}^n \times [t,T] \times \Omega$. We further suppose that $D^2 v_\delta^\epsilon$ is uniformly bounded in ϵ for fixed δ. ◀

In many applications, the smoothing process of section V.7 in [FS06] may be used to built semiconcave/convex sub/supersolutions.

To obtain a non-random Effective Hamiltonian, one must impose certain ergodicity conditions on the data of the problem. So far, we have made no assumptions on the behavior of L, f or σ with respect to the random parameter ω. It turns out that a convenient hypothesis is stationarity, as we put forward next. For the purposes of section 7, the next assumption is enough, but later we will also require certain ergodicity hypothesis to obtain a non-random homogenized limit, see Assumption 13 below.

Assumption 7. We assume that there is an action τ of $\mathbb{R}^n$ in the probability space Ω, denoted by $\tau_X\omega$, for $X \in \mathbb{R}^n$, which preserves the measure in Ω. Furthermore, we suppose that the L, f and σ are stationary with respect to this action, that is, for all $Y \in \mathbb{R}^n$

$$L(x, X+Y, v, \omega) = L(x, X, v, \tau_Y\omega), \tag{11}$$

$$f(x, X+Y, v, \omega) = f(x, X, v, \tau_Y\omega), \tag{12}$$

and

$$\sigma(x, X+Y, \epsilon, v, \omega) = \sigma(x, X, \epsilon, v, \tau_Y\omega). \tag{13}$$

◀

As an example, note that the periodic setting can be embedded in this framework in the following way: we let $\Omega = \mathbb{T}^n$, and consider in Ω the (normalized) Lebesgue measure. Then, if $\hat{L}(x,X,v)$, $\hat{f}(x,X,v)$ and $\hat{\sigma}(x,X,v)$ are periodic in X, we can write

$$L(x,X,v,\omega) = \hat{L}(x, X+\omega, v),$$

with similar definitions for f and σ, and then we choose $\tau_Y\omega = \omega + Y$.

Next, we state the concept of tight sequence of measures, by generalizing a standard definition in probability theory (see, e.g., page 336 of [Bil95]):

Definition 1. *A net of probability measures μ_α on a given space $Z \times \Omega$, with $Z \subset \mathbb{R}^d$ is said to be tight if for any $\epsilon > 0$ there exists a compact set K_ϵ in such a way that*

$$\limsup_\alpha \mu_\alpha\Big((Z \setminus K_\epsilon) \times \Omega\Big) \leq \epsilon\,.$$

The following assumption is simply a compactness property on measures on Ω that is necessary to extract weak limits.

Assumption 8. We assume that for each tight net of probability measures $\eta_\alpha(z,\omega)$ on $Z\times\Omega$, with $Z\subset\mathbb{R}^d$, there exists a subnet which converges to a probability measure $\eta(z,\omega)$, in the sense that: if $\psi(X,z,\omega):\mathbb{R}^d\times Z\times\Omega\to\mathbb{R}$ is continuous in (X,z) and stationary (that is, $\psi(X+Y,z,\omega)=\psi(X,z,\tau_Y\omega)$ for any $(X,Y,z,\omega)\in\mathbb{R}^d\times\mathbb{R}^d\times Z\times\Omega$), we have that

$$\lim_\alpha \int_{Z\times\Omega}\psi(z,0,\omega)\,d\eta_\alpha(z,\omega) \;=\; \int_{Z\times\Omega}\psi(z,0,\omega)\,d\eta(z,\omega)\,,$$

up to subnets. ◀

In section 7 we will use the following scaling hypothesis for the diffusion coefficient σ

Assumption 9. We will assume that the viscosity coefficient σ is either

A. $\sigma=\epsilon\sigma_0$
B. $\sigma=\epsilon^{1/2}\sigma_0$
C. $\sigma=\sigma_0$,

with $\sigma_0=\sigma_0(x',X,v,\omega):\mathbb{R}^n\times\mathbb{R}^n\times U\times\Omega\to\mathrm{Mat}(n\times m)$. ◀

We should remark that these are the relevant scalings, as many others could be studied with similar methods.
From section 9 on, we will assume that our viscosity solutions are smooth, up to a small uniformly elliptic correction of the Hamiltonian. That is, we suppose:

Assumption 10. For any $\ell\in\mathbb{N}$, there exist $u_\ell^\epsilon\in C^2(\mathbb{R}^n\times(0,T))$, $\psi_\ell\in C^2(\mathbb{R}^n)$ and a Hamiltonian H_ℓ such that u_ℓ^ϵ is a solution of the following Hamilton-Jacobi equation:

$$\begin{cases} -\,\partial_t u_\ell^\epsilon + H_\ell\left(D_x^2u_\ell^\epsilon, D_xu_\ell^\epsilon, x, \dfrac{x}{\epsilon}, \epsilon, \omega\right)=0\\ u_\ell^\epsilon(x,T,\omega) \;=\; \psi_\ell(x)\,, \end{cases}$$

and in such a way that

$$\begin{aligned} &\lim_{\ell\to+\infty} u_\ell^\epsilon(x,t)=u^\epsilon(x,t)\\ &\lim_{\ell\to+\infty}\psi_\ell(x)=\psi(x)\,,\\ \text{and}\qquad &\lim_{\ell\to+\infty}\sup_x\Big|H_\ell\Big(D_x^2\Psi(x),D_x\Psi(x),x,\frac{x}{\epsilon},\epsilon,\omega\Big)\\ &\qquad\qquad -H\Big(D_x^2\Psi(x),D_x\Psi(x),x,\frac{x}{\epsilon},\epsilon,\omega\Big)\Big|=0 \end{aligned}$$

for any $x\in\mathbb{R}^n$, $t\in[0,T]$ and $\Psi\in C^2(\mathbb{R}^n)$ with bounded C^2-norm. Also, $\overline{H}_\ell$ satisfies

$$-\,D_{M_{ij}}H_\ell\,\xi_i\,\xi_j\;\geq\;c_\ell\,|\xi|^2\,, \tag{14}$$

for some $c_\ell>0$.
We also assume that

$$\sup_{x\in\mathbb{R}^n,\,t\in[0,T]}|Du_\ell^\epsilon|\;<\;+\infty\,. \tag{15}$$

◀

A natural possibility for such Hamiltonian H_ℓ is of course $H - \operatorname{tr} M/\ell$, but other regularizations are possible (for general regularization results, see, e.g., [Eva82, Eva97, Wan90, Wan92a, Wan92b, Wan92c, CC95, LS05]).
We remark that (15) is not assumed, a-priori, to be uniform in ϵ and ℓ (in fact, uniform estimates will be proved in Proposition 20 below). In many applications, bounds as in (15) are quite natural and may be obtained by comparing trajectories (see, e.g., page 555 of [Eva98] and Lemma 8.1 of [FS06]).
The next three assumptions will be used in section 9 to prove various uniform estimates in ϵ.

Assumption 11. We assume that H enjoys one of the following properties:

P1) $H(M,p,x,X,\epsilon,\omega)$ does not depend on M and satisfies the coercivity hypothesis

$$\lim_{|p|\to\infty} H(p,x,X,\epsilon,\omega) = +\infty, \tag{16}$$

uniformly in $(x,X,\epsilon,\omega) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^+ \times \Omega$.

P2) There exists a function $F(p,\zeta)$ such that for all $(M,p,x,X,\epsilon,\omega) \in \operatorname{Mat}(n\times n)\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^+\times\Omega$ and $\zeta\in\mathbb{R}$ such that

$$\zeta + H(M,p,x,X,\epsilon,\omega) = 0$$

implies

$$-\frac{1}{\epsilon}D_{M_{jk}}H\Big(\frac{1}{\epsilon}M,p,x,X,\epsilon,\omega\Big)M_{ij}M_{ik} + \epsilon p_i D_{x_i}H\Big(\frac{1}{\epsilon}M,p,x,X,\epsilon,\omega\Big)$$
$$+p_i D_{X_i}H\Big(\frac{1}{\epsilon}M,p,x,X,\epsilon,\omega\Big) \ \geq\ F(p,\zeta),$$

and, furthermore, if K is a fixed compact set of $\mathbb{R}^n$,

$$\inf_{\zeta\in K} F(p,\zeta) \to +\infty, \tag{17}$$

as $|p|\to+\infty$.

◀

Examples of Hamiltonians satisfying P2 are easily obtained, for instance, by considering the case $\sigma := \mathrm{Id}_n$, $f := v$, $L := |v|^2/2 - U(x,X,\omega)$. In such a case, (6) yields that

$$H = -\frac{1}{2}\operatorname{tr} M + \frac{1}{2}|p|^2 + U(x,X,\omega)\,, \tag{18}$$

which is plainly seen to satisfy P2.
The next assumption will be used in section 9 to prove that the limit is non-random.

Assumption 12. We suppose that there exists a constant C independent of ϵ such that for any $(x,z)\in\mathbb{R}^n\times\mathbb{R}^n$ we have

$$\left|D_x H\left(D^2u^\epsilon(z,t), Du^\epsilon(z,t), z, \frac{x}{\epsilon},\epsilon,\omega\right)\right| \leq C.$$

◀

Note that (18) provides an example of an Hamiltonian satisfying Assumption 12, due to Proposition 20 below.

Assumption 13. We assume the following uniform ergodicity hypothesis: for any $\delta > 0$ there exists $M > 0$, such that for any ω_0, $\omega_1 \in \Omega$, $x \in \mathbb{R}^n$, $t \in [0,T]$, there exists $y \in \mathbb{R}^n$, $|y| \le M$, such that

$$\Big|H\Big(D^2u^\epsilon(x,t,\omega_1), Du^\epsilon(x,t,\omega_1), x, \frac{x}{\epsilon}, \epsilon, \tau_y\omega_0\Big) \\ -H\Big(D^2u^\epsilon(x,t,\omega_1), Du^\epsilon(x,t,\omega_1), x, \frac{x}{\epsilon}, \epsilon, \omega_1\Big)\Big| \le \delta.$$

◀

Heuristically, Assumption 13 means we are be able to get close to any point of Ω from any other by translating, the size of such translation depending only on how close we want to get (this is the case, for instance, of the geodesics with irrational slope on the flat torus).
From section 11 on, we will also suppose the following:

Assumption 14. We assume that there are constants $\bar{C}_0$, $\bar{C}_1$ and $\bar{C}_2$ such that

$$|L(x', X, 0, \omega)| \le \bar{C}_0, \qquad |f(x', X, 0, \omega)| \le \bar{C}_1, \qquad |\sigma(x', X, 0, \omega)| \le \bar{C}_2.$$

for any $(x', X, \omega) \in \mathbb{R}^n \times \mathbb{R}^n \times \Omega$. ◀

Assumption 15. Let $p > 1$ be as in Assumption 4. We suppose that

$$|f(x', X, v, \omega)|^2 \le C(1 + L(x', X, v, \omega)), \qquad |\sigma(x', X, \epsilon, v, \omega)| \le C$$

and

$$|\sigma\sigma^T(x', X, \epsilon, v, \omega)|^p \le C(1 + L(x', X, v, \omega)),$$

for any $(x', X, v, \omega) \in \mathbb{R}^n \times \mathbb{R}^n \times U \times \Omega$. ◀

We now deal with the proof of Theorem 1.

4. Optimal control and linear programming

In this section we are going to convert the original optimal control problem (1) into a linear programming problem. As the random parameter ω plays no role in most of this discussion we omit it until section 7.
The controlled stochastic dynamics (3) has an infinitesimal generator (see, e.g., page 105 in [Eva06b]), given by

$$A\varphi(x) = f \cdot D_x\varphi + \frac{\sigma\sigma^T}{2} : D_x^2\varphi. \tag{19}$$

To encode both the behavior at the slow scale x and fast scale $X = \frac{x}{\epsilon}$, we need to introduce the extended generator $\tilde{A}$, given by

$$\tilde{A}\phi(x, X) = f \cdot D_x\phi + \frac{\sigma\sigma^T}{2} : D_x^2\phi \\ + \frac{1}{\epsilon} f \cdot D_X\phi + \frac{1}{\epsilon}(\sigma\sigma^T) : D_{xX}^2\phi + \frac{1}{\epsilon^2}\frac{\sigma\sigma^T}{2} : D_X^2\phi.$$

This extended generator satisfies the following identity

$$(\tilde{A}\phi)\left(x, \frac{x}{\epsilon}\right) = A\left(\phi\left(x, \frac{x}{\epsilon}\right)\right) .$$

Observe further that

$$\tilde{A}\left(\varphi\left(X - \frac{x}{\epsilon}\right)\right) = 0 \,, \tag{20}$$

for all C^2 functions $\varphi(x) : \mathbb{R}^n \to \mathbb{R}$.

Before proceeding, we need some further definitions. If U is bounded, set $\gamma = 1$, otherwise, let $\gamma : U \to \mathbb{R}^+$ be a function which satisfies

$$\lim_{|v|\to\infty} \inf_{x,X} \frac{L(x, X, v, \omega)}{\gamma(v)} = +\infty,$$

and

$$\lim_{|v|\to\infty} \frac{\tilde{A}^v \phi(x, X)}{\gamma(v)} = 0,$$

for all $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n)$. Let p be the exponent for which Assumption 4 holds. Consider the spaces $\mathcal{P}$ and $\mathcal{Q}$ of signed[1] measures, resp. μ on $\mathbb{R}^n \times \mathbb{R}^n \times U \times [t, T]$ and ν on $\mathbb{R}^n \times \mathbb{R}^n$, which satisfy

$$\int \gamma(v) + |x|^p + \epsilon|X|^p d|\mu| < \infty, \qquad \int |x|^p + \epsilon|X|^p d|\nu| < \infty.$$

We remark that $\mathcal{P}$ is the dual of the set $C_*(\mathbb{R}^n \times \mathbb{R}^n \times U \times [t, T])$ of continuous functions $\chi : \mathbb{R}^n \times \mathbb{R}^n \times U \times [t, T] \to \mathbb{R}$, which satisfy

$$\lim_{|x|+|X|+|v|\to\infty} \sup_{t'} \frac{\chi(x, X, v, t')}{\gamma(v) + |x|^p + \epsilon|X|^p} = 0,$$

and $\mathcal{Q}$ is the dual of the set $C_\sharp(\mathbb{R}^n \times \mathbb{R}^n)$ of continuous functions on $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, which satisfy

$$\lim_{|x|+|X|\to\infty} \frac{\eta(x, X)}{|x|^p + \epsilon|X|^p} = 0.$$

We further consider the space $C_\star(\mathbb{R}^n \times \mathbb{R}^n \times [t, T])$ of continuous functions on $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, which satisfy

$$\lim_{|x|+|X|\to\infty} \sup_{t'} \frac{\phi(x, X, t')}{|x|^p + \epsilon|X|^p} = 0.$$

In this section, we relax the original stochastic control problem by looking at generalized trajectories represented by certain measures, which will be elements of $\mathcal{P}$ and $\mathcal{Q}$. To do this, we must consider measures that satisfy a generalized holonomy

[1]All the measures will be tacitly assumed to be Radon. If not differently specified, they are also assumed to be positive.

constraint. More precisely, let ν_t^ϵ be a given probability measure. We look for measures $(\mu^\epsilon(x', X, v, s), \nu_T^\epsilon(x', X)) \in \mathcal{P} \times \mathcal{Q}$ which satisfy

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} & \partial_t\phi(x', X, s) + \tilde{A}\phi(x', X, s)\, d\mu^\epsilon(x', X, v, s; \omega) \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x', X, T\right)\, d\nu_T^\epsilon(x', X) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x', X, t\right)\, d\nu_t^\epsilon(x', X)\,,\end{aligned} \tag{21}$$

for any $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n \times [t, T])$ such that $\partial_t\phi(x', X, s) + \tilde{A}\phi(x', X, s) \in C_*(\mathbb{R}^n \times \mathbb{R}^n \times U \times [t, T])$ and $\phi\left(x', X, t\right), \phi\left(x', X, T\right) \in C_\sharp(\mathbb{R}^n \times \mathbb{R}^n)$.
The measure ν_t^ϵ encodes the given initial condition $\mathbf{x}(t) = x$, allowing for probability distributions of initial conditions, whereas the measure ν_T^ϵ, which is arbitrary, encodes the possible probability distribution of terminal conditions. We will always choose

$$\nu_t^\epsilon = \delta_x(x')\delta_{\frac{x}{\epsilon}}(X) \tag{22}$$

and we will omit the dependence of ν_t^ϵ and ν_T^ϵ upon the prescribed initial point x.
The motivation for this holonomy property is the following: let us consider a trajectory $(\mathbf{x}(\cdot), \mathbf{v}(\cdot))$ of the controlled dynamics (3); to this trajectory, one can associate measures μ and ν_T in the following way

$$\begin{aligned}&\int \phi(x', X, v, s)\, d\mu^\epsilon(x', X, v, s; \omega) \\ &:= E\int_t^T \phi\left(\mathbf{x}(s), \frac{\mathbf{x}(s)}{\epsilon}, \mathbf{v}(s), s\right)\, ds\,,\end{aligned} \tag{23}$$

and

$$\int \phi(x', X)\, d\nu_T^\epsilon(x', X) := E\phi\left(\mathbf{x}(T), \frac{\mathbf{x}(T)}{\epsilon}\right). \tag{24}$$

By Dynkin's Formula (see, e.g., formula (3) on page 105 of [Eva06b]), the constraint in (21) is satisfied by these measures.
Of course, by considering measures instead of trajectories, we are, in principle, enlarging the minimizing set. However, as we will see later, the value of the infimum is unchanged (see Propositions 4 and 9 below for further details on this).
In the next proposition, we show that a measure μ satisfying (21) has always total mass $T - t$.

Proposition 2. *Suppose that μ is a measure on $\mathbb{R}^n\times\mathbb{R}^n\times U\times[t, T]$ which satisfies (21), for a suitable probability measure ν_T^ϵ on $\mathbb{R}^n \times \mathbb{R}^n$. Then*

$$\int v(s)d\mu = \int_t^T v(s)ds \tag{25}$$

for any $v \in C([t, T])$. Also,

$$\mu(\mathbb{R}^n \times \mathbb{R}^n \times U \times [t, T]) \,=\, T - t. \tag{26}$$

Proof. Choose $\Upsilon(t)$ such that $\upsilon(t) = \Upsilon'(t)$. Then, from (21), we have that

$$\int \upsilon(s)d\mu = \int \Upsilon(T)d\nu_T^\epsilon - \int \Upsilon(t)d\nu_t^\epsilon = \Upsilon(T) - \Upsilon(t) = \int_t^T \upsilon(s)ds,$$

yielding (25). Then (26) follows by applying (25) to $\upsilon := 1$. □

We now prove that both ν_T^ϵ and μ^ϵ are supported on $\{X = \frac{x'}{\epsilon}\}$, as long as so is ν_t^ϵ.

Proposition 3. *Suppose ν_t^ϵ is supported in the set $X = \frac{x'}{\epsilon}$, and that $(\mu^\epsilon, \nu_T^\epsilon)$ satisfy* (21)*. Then both ν_T^ϵ and μ^ϵ are supported in the set $X = \frac{x'}{\epsilon}$.*

Proof. Let $\varkappa : \mathbb{R}^n \to \mathbb{R}_0^+$ be a C^2 function such that $\varkappa(0) = 0$ and $\varkappa(x') > 0$ if $x' \neq 0$. Assume that $\varkappa\left(X - \frac{x'}{\epsilon}\right) \in C_\sharp(\mathbb{R}^n \times \mathbb{R}^n)$. To prove the desired claim, we will show that if

$$\int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\nu_t^\epsilon(x', X) = 0,$$

then

$$\int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\nu_T^\epsilon(x', X) = 0,$$

and

$$\int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\mu^\epsilon(x', X, v, s; \omega) = 0.$$

Set $\phi(x', X) := \varkappa\left(X - \frac{x'}{\epsilon}\right)$. Since, by (20),

$$\tilde{A}\phi = 0\,, \tag{27}$$

the constraint in (21) implies that

$$\int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\nu_T^\epsilon = \int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\nu_t^\epsilon = 0. \tag{28}$$

Now choose $\phi(x', X, t) := t\varkappa\left(X - \frac{x'}{\epsilon}\right)$. Then, using (27), (28) and (21) once more, we gather that

$$\int \varkappa\left(X - \frac{x'}{\epsilon}\right) d\mu^\epsilon = \int \partial_t\phi + \tilde{A}\phi\, d\mu^\epsilon = 0$$

as we wished. □

Proposition 4. *For $t \in (0, T)$, let*

$$\begin{aligned} \tilde{u}^\epsilon(x, t, \omega) \;\; &:= \;\; \inf \int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} L(x', X, v, \omega)\, d\mu^\epsilon(x', X, v, s; \omega) \\ &\qquad + \int_{\mathbb{R}^n\times\mathbb{R}^n} \psi(x')\, d\nu_T^\epsilon(x', X; \omega)\,, \end{aligned} \tag{29}$$

where the infimum is taken over all measures $(\mu^\epsilon, \nu_T^\epsilon)$ *which, for every* $\omega \in \Omega$*, satisfy the following:* μ^ϵ *is a measure on* $\mathbb{R}^n \times \mathbb{R}^n \times U \times [t,T]$ *with total mass* $T-t$*,* ν_T^ϵ *is a probability measure on* $\mathbb{R}^n \times \mathbb{R}^n$*, and* μ^ϵ *and* ν_T^ϵ *satisfy the following constraint:*

$$\begin{aligned}(30)\qquad &\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} \partial_t\phi(x',X,s) + \tilde{A}\phi(x',X,s)\,d\mu^\epsilon(x',X,v,s;\omega)\\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x',X,T\right)\, d\nu_T^\epsilon(x',X;\omega) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi\left(x',X,t\right)\, d\nu_t^\epsilon(x',X)\,,\end{aligned}$$

for any $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n \times [t,T])$*. Then,* $u^\epsilon \geq \tilde{u}^\epsilon$*.*

Proof. Let u^ϵ be as in (5) and $\tilde{u}^\epsilon$ given by (29). Fix $\delta > 0$, and consider a bounded progressively measurable control $\mathbf{v}_\delta$, and the corresponding trajectory $\mathbf{x}_\delta$ of (3) with the initial condition $\mathbf{x}_\delta(t) = x$, such that

$$\begin{aligned}(31)\qquad &u^\epsilon(x,t,\omega)+\delta\\ &\geq E\left[\int_t^T L\left(\mathbf{x}_\delta(s), \frac{\mathbf{x}_\delta(s)}{\epsilon}, \mathbf{v}_\delta(s), \omega\right) ds + \psi\left(\mathbf{x}_\delta(T)\right)\right].\end{aligned}$$

We introduce the probability measures μ_δ and ν_δ on $\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]$ and on $\mathbb{R}^n\times\mathbb{R}^n$, respectively, defined by

$$\begin{aligned}&\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} \phi(x',X,v,s)\,d\mu_\delta(x',X,v,s)\\ &\qquad := E\int_t^T \phi\left(\mathbf{x}_\delta(s), \frac{\mathbf{x}_\delta(s)}{\epsilon}, \mathbf{v}_\delta(s), s\right) ds\end{aligned}$$

and

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi(x',X)\,d\nu_\delta(x',X) := E\varphi\left(\mathbf{x}_\delta(T), \frac{\mathbf{x}_\delta(T)}{\epsilon}\right),$$

for any $\phi \in C(\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T])$ and $\varphi \in C(\mathbb{R}^n\times\mathbb{R}^n)$.
Then, if $\phi \in C^2(\mathbb{R}^n\times\mathbb{R}^n\times[t,T])$, by Dynkin's Formula (see, e.g., formula (3) on page 105 of [Eva06b]),

$$\begin{aligned}&\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]} \partial_t\phi + \tilde{A}\phi d\mu_\delta\\ &= E\left[\phi\left(\mathbf{x}_\delta(T), \frac{\mathbf{x}_\delta(T)}{\epsilon}, T\right) - \phi\left(\mathbf{x}_\delta(t), \frac{\mathbf{x}_\delta(t)}{\epsilon}, t\right)\right]\\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x',X,T)d\nu_\delta(x',X) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x',X,t)d\nu_t^\epsilon(x',X),\end{aligned}$$

where ν_t^ϵ is taken as

$$\int \phi(x',X,t)\,d\nu_t^\epsilon(x',X) := \phi\left(x, \frac{x}{\epsilon}, t\right),$$

which shows that μ_δ and ν_δ satisfy the constraint in (30). Accordingly, (29) and (31) give that $u^\epsilon + \delta \geq \tilde{u}^\epsilon$, and so the claim follows by sending δ to zero. ☐

5. Moment estimates

In this section, we discuss some a-priori moment estimates for minimizing measures $(\mu^\epsilon, \nu_T^\epsilon)$, as given in (29), that will be essential to establish the tightness necessary to our scopes. More refined estimates, needed to prove that the homogenized limit is a viscosity solution of a Hamilton-Jacobi equation, will be developed in section 11. From this section on, we will assume Assumptions 4 and 5. By comparison with the control $\mathbf{v} := 0$ in (3), recalling Assumption 5, we see that

$$\int L(x', X, v, \omega) d\mu^\epsilon \le C\,, \tag{32}$$

with $C > 0$ independent of ϵ.

Proposition 5. *Let $(\mu^\epsilon, \nu_T^\epsilon)$ be measures satisfying (30), with ν_t^ϵ given by (22). Let L satisfy (32). Let p be as in Assumption 4. Then, there exists $C > 0$, possibly depending on t and T, but independent of ϵ, in such a way that*

$$\int |x'|^p \, d\mu^\epsilon + \int |x'|^p \, d\nu_T^\epsilon \le C\,(1 + |x|^p)\,.$$

Proof. In what follows, we will denote by C_i suitable positive quantities, which may depend on t, T, but that are independent of ϵ.
We observe that

$$|x'|^p \le (1 + |x'|^2)^{p/2} \le 2^{p/2}(1 + |x'|^p)\,. \tag{33}$$

We apply the constraint in (30) to the function $\phi(x') := (1 + |x'|^2)^{p/2}$ and we obtain that

$$\begin{aligned} &\int p(1 + |x'|^2)^{(p-2)/2} f \cdot x' \\ &+\frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^T)_{ij}\Big(p(p-2)(1 + |x'|^2)^{(p-4)/2} x'_i x'_j + p(1 + |x'|^2)^{(p-2)/2}\delta_{ij}\Big)\, d\mu^\epsilon \\ &= \int (1 + |x'|^2)^{p/2}\, d\nu_T^\epsilon - \int (1 + |x'|^2)^{p/2}\, d\nu_t^\epsilon, \end{aligned}$$

and so, by (33),

$$\begin{aligned} \int |x'|^p d\nu_T^\epsilon \;&\le\; \int (1 + |x'|^2)^{p/2} d\nu_T^\epsilon \\ &\le\; C_2\Big(\int |f|\,(1 + |x'|^2)^{(p-1)/2} \\ &\quad + |\sigma\sigma^T|\,(1 + |x'|^2)^{(p-2)/2} d\mu^\epsilon + \int (1 + |x'|^2)^{p/2} d\nu_t^\epsilon\Big)\,. \end{aligned}$$

Consequently, Young's inequality, Assumption 4, estimates (32) and (33), and Proposition 2 give that

$$
\begin{aligned}
&\int |x'|^p d\nu_T^\epsilon \\
&\leq C_3 \left(\int |f|^p + \frac{|\sigma\sigma^T|^p}{(1+|x'|^2)^{p/2}} + (1+|x'|^2)^{p/2}\, d\mu^\epsilon + \int (1+|x'|^2)^{p/2} d\nu_t^\epsilon \right) \\
&\leq C_4 \left(\int 1 + L + |x'|^p\, d\mu^\epsilon + \int 1 + |x'|^p d\nu_t^\epsilon \right) \\
&\leq C_5 \left(1 + \int |x'|^p\, d\mu^\epsilon + \int |x'|^p d\nu_t^\epsilon \right) .
\end{aligned} \tag{34}
$$

To estimate this last term, we apply again (21) to the function $\phi(x', s) := (s-T)(1+|x'|^2)^{p/2}$, to obtain

$$
\begin{aligned}
&\int (1+|x'|^2)^{p/2} + (s-T)\Big[p(1+|x'|^2)^{(p-2)/2} f\cdot x' \\
&+\frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^T)_{ij} \Big(p(p-2)(1+|x'|^2)^{(p-4)/2} x_i' x_j' + p(1+|x'|^2)^{(p-2)/2}\delta_{ij} \Big) \Big]\, d\mu^\epsilon \\
= \ & (T-t) \int (1+|x'|^2)^{p/2} d\nu_t^\epsilon.
\end{aligned}
$$

Thus, using (33) and a scaled Young's inequality, we obtain that

$$
\begin{aligned}
&\int (1+|x'|^2)^{p/2}\, d\mu^\epsilon \\
\leq \ & C_6 \left(1 + \int |x'|^p\, d\nu_t^\epsilon + \int |f|^p + \frac{|\sigma\sigma^T|^p}{(1+|x'|^2)^{p/2}}\, d\mu^\epsilon \right) \\
&+\frac{1}{2}\int (1+|x'|^2)^{p/2}\, d\mu^\epsilon .
\end{aligned}
$$

Thus, exploiting Assumption 4, estimate (32) and Proposition 2, we conclude that

$$
\begin{aligned}
\frac{1}{2}\int |x'|^p d\mu^\epsilon \ \leq \ & \frac{1}{2}\int (1+|x'|^2)^{p/2} d\mu^\epsilon \\
\leq \ & C_7 \left(\int 1 + L d\mu + \int |x'|^p\, d\nu_t^\epsilon \right) \\
\leq \ & C_8 \left(1 + \int |x'|^p\, d\nu_t^\epsilon \right) .
\end{aligned}
$$

This, (34) and the definition of ν_t^ϵ, see (22), yield the desired result. $\square$

As a consequence of this result, using (22), we have, using Proposition 3, that any minimizing sequence $(\mu_n^\epsilon, \nu_{T,n}^\epsilon)$ will satisfy the bound

$$
\int |x'|^p + \epsilon^p |X|^p d\mu_n^\epsilon \leq C, \qquad \int |x'|^p + \epsilon^p |X|^p d\nu_{T,n}^\epsilon \leq C. \tag{35}
$$

The above tightness and Assumption 8 imply that there are minimizing measures μ^ϵ and ν_T^ϵ attaining the infimum in (29).

6. Duality

In addition to the Hamiltonian H as defined in (6), we need also the extended Hamiltonian operator

$$\tilde{H}(D^2\phi(x,X), D\phi(x,X), x, X, \epsilon, \omega) := \sup_{v\in U} -\tilde{A}\phi - L, \tag{36}$$

which is defined for $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n)$. It will be useful later the observation that

$$\tilde{H}(D^2\phi_\lambda(x,X), D\phi_\lambda(x,X), x, X, \epsilon, \omega),$$

with

$$\phi_\lambda = \phi(x,X) + \lambda\varphi(X - \frac{x}{\epsilon}), \tag{37}$$

is constant in λ, for any C^2 function φ, thanks to (20).
Recalling the framework introduced on page 10, we define the set C^2_Λ to be the set of functions $\phi \in C^2(\mathbb{R}^n \times \mathbb{R}^n \times [t,T])$ such that $\phi(x',X,t), \phi(x',X,T) \in C_\sharp(\mathbb{R}^n \times \mathbb{R}^n)$, and $\partial_t\phi + \tilde{A}^v\phi \in C_*(\mathbb{R}^n \times \mathbb{R}^n \times U \times [t,T])$.
In order to prove the reverse inequality of the one in Proposition 4, we introduce the dual problem of (29):

Proposition 6. *In the notation of Proposition 4, we have that*

$$\begin{aligned} \tilde{u}^\epsilon(x,t,\omega) = & \\ & \sup_{\phi(x',X,s)} (T-t)\left[\inf_{(x',X,s)\in\mathbb{R}^n\times\mathbb{R}^n\times(t,T)} \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right)\right] \\ & + \inf_{x',X}\left[\psi(x',X,\omega) - \phi(x',X,T)\right] + \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x',X,t)\, d\nu_t^\epsilon(x',X). \end{aligned} \tag{38}$$

where the supremum is taken over any $\phi \in C^2_\Lambda$.

To heuristically motivate the statement in (38), we apply the minimax principle, that is, we exchange formally the infimum with the supremum in the variational problem (29), by using the constraint in (30), thus obtaining:

$$\begin{aligned}
&\tilde{u}^\epsilon(x,t,\omega)\\
= &\inf_{\mu^\epsilon,\nu^\epsilon_T}\sup_{\phi(x',X,t)}\left[\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times\Omega\times(t,T)} L(x',X,v,\omega)+\tilde{A}\phi+\partial_t\phi\,d\mu^\epsilon\right.\\
&\left.+\int_{\mathbb{R}^n\times\mathbb{R}^n\times\Omega}(\psi-\phi)d\nu^\epsilon_T+\int_{\mathbb{R}^n\times\mathbb{R}^n\times\Omega}\phi d\nu^\epsilon_t\right]\\
&\text{(using the minimax principle)}\\
= &\sup_{\phi(x',X,s)}\left[\int_{\mathbb{R}^n\times\mathbb{R}^n\times\Omega}\phi d\nu^\epsilon_t\right.\\
&+\inf_{\mu^\epsilon}\int_{\mathbb{R}^n\times\mathbb{R}^n\times U\times\Omega\times(t,T)} L(x',X,v,\omega)+\tilde{A}\phi+\partial_t\phi\,d\mu^\epsilon\\
&\left.+\inf_{\nu^\epsilon_T}\int_{\mathbb{R}^n\times\mathbb{R}^n\times\Omega}(\psi-\phi)d\nu^\epsilon_T\right],
\end{aligned}$$

and then, taking into account that both μ^ϵ and ν^ϵ_T are measures, with ν^ϵ_T a probability measure, and μ^ϵ with mass $T-t$, the definition of $\tilde{A}$ and $\tilde{H}$ we obtain (38).

Proof (of Proposition 6). The rigorous proof of Proposition 6 will be quite long (ending on page 20) and it will make use of the Legendre-Fenchel-Rockafellar Theorem. Let $\mathcal{P}$ and $\mathcal{Q}$ be defined as in page 10. Define the following sets:

$$\mathcal{M}_1 := \left\{(\mu,\nu)\in\mathcal{P}\times\mathcal{Q} : \int d\mu = T-t,\ \int d\nu = 1,\ \mu,\nu\geq 0\right\},$$

and

$$\mathcal{M}_2 := \left\{(\mu,\nu)\in\mathcal{P}\times\mathcal{Q} : \int \tilde{A}\phi+\partial_t\phi d\mu = \int\phi d\nu - \int \phi d\nu^\epsilon_t, \forall\phi\in C^2_\Lambda\right\}.$$

Given $\chi\in C_*(\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T])$ and $\phi\in C_\star(\mathbb{R}^n\times\mathbb{R}^n\times[t,T])$, we set

$$\begin{aligned}
h(\chi,\phi) \;:=\;& (T-t)\sup_{(x,X,v,s)}\left[-\chi(x,X,v,s)-L(x,X,v,s,\omega)\right]\\
&+\sup_{(x',X)}\left[\psi(x',X,\omega)-\phi(x',X,T)\right].
\end{aligned}\tag{39}$$

Note that, for short, we have omitted the dependence of h on ω. We consider the set

$$\mathcal{C} = \mathrm{cl}\left\{(\chi,\phi)\ :\ \phi\in C^2_\Lambda\ ,\ \chi=\tilde{A}\phi+\partial_t\phi\in C_*\right\}\ ,$$

where the closure is taken in the $C^2_\Lambda\times C_*$ topology.
Let $(\mu_*,\nu_*)\in\mathcal{M}_1\cap\mathcal{M}_2$ and define

$$g(\chi,\phi) := \begin{cases}\int\chi d\mu_* - \int\phi(x,X,T)d\nu_* & \text{if } (\chi,\phi)\in\mathcal{C}\\ -\infty & \text{otherwise.}\end{cases}$$

We will now compute the dual functions of h and g, that is

$$\begin{aligned} h^*(\mu,\nu) &:= \sup_{(\chi,\phi)\in C_*\times C_\star} -\int \chi d\mu - \int \phi d\nu - h(\chi,\phi)\,, \\ g^*(\mu,\nu) &:= \inf_{(\chi,\phi)\in C_*\times C_\star} -\int \chi d\mu - \int \phi d\nu - g(\chi,\phi)\,. \end{aligned} \tag{40}$$

Lemma 7. *We have*

$$h^*(\mu,\nu) = \begin{cases} \int L d\mu + \int \psi d\nu & \text{if } (\mu,\nu)\in\mathcal{M}_1 \\ +\infty & \text{otherwise} \end{cases} \tag{41}$$

and

$$g^*(\mu,\nu) = \begin{cases} 0 & \text{if } (\mu,\nu)\in\mathcal{M}_2 \\ -\infty & \text{otherwise.} \end{cases} \tag{42}$$

Proof. We first claim that if either μ or ν are non-positive then $h^*(\mu,\nu)=+\infty$. To see this, we choose a sequence of non-negative functions (χ_j,ϕ_j) such that

$$-\int \chi_j d\mu - \int \phi_j d\nu \to +\infty.$$

Since $L\geq 0$ we have that

$$\sup -\chi_j - L \leq 0,$$

and, furthermore,

$$\sup \psi - \phi_j \leq \sup \psi\,.$$

As a consequence, since ψ is bounded,

$$h(\chi_j,\phi_j) \leq -(T-t)\inf L + \sup\psi$$

is bounded from above uniformly in j, and so $h^*(\mu,\nu)=+\infty$.
Now we claim that the following inequality holds:

$$\begin{aligned} h^*(\mu,\nu) \geq & \int L d\mu + \int \psi d\nu \\ & + \sup_{\tilde\chi}\left[\int \tilde\chi d\mu - (T-t)\sup\tilde\chi\right] + \sup_{\tilde\phi}\left[\int\tilde\phi d\nu - \sup\tilde\phi\right]. \end{aligned} \tag{43}$$

To establish (43), consider a sequence $L_j \geq \inf L = 0$ (see Assumption 1) of compactly supported functions, increasing pointwise to L. For any $\chi\in C_*$ and $\phi\in C_\star$, define $\tilde\chi$ and $\tilde\phi$ by $\chi=-L_j-\tilde\chi$, and $\phi=-\psi-\tilde\phi$. Note that $\tilde\chi\in C_*$ and $\tilde\phi\in C_\star$, so that

$$\begin{aligned} h^*(\mu,\nu) \geq \sup_{\tilde\chi,\tilde\phi} \Bigg[& \int L_j d\mu + \int \tilde\chi d\mu + \int \psi d\nu + \int \tilde\phi d\nu \\ & -(T-t)\sup\left[L_j+\tilde\chi-L\right] - \sup\tilde\phi\Bigg]. \end{aligned}$$

Since $L_j - L \leq 0$ we gather that

$$h^*(\mu,\nu) \geq \int L_j d\mu + \int \psi d\nu$$
$$+ \sup_{\tilde{\chi}} \left[\int \tilde{\chi} d\mu - (T-t) \sup \tilde{\chi} \right] + \sup_{\tilde{\phi}} \left[\int \tilde{\phi} d\nu - \sup \tilde{\phi} \right].$$

By letting $j \to +\infty$ we obtain (43), due to the Monotone Convergence Theorem.
Now we claim that if the mass of μ is not $T-t$ or the mass of ν is not 1, then $h^*(\mu,\nu) = +\infty$. To see this just add constants to $\tilde{\chi}$ or to $\tilde{\phi}$.
Finally, if the mass of μ is $T-t$ and the mass of ν is 1, by choosing $\tilde{\chi} := 0$ and $\tilde{\phi} := 0$ in (43), we obtain

$$h^*(\mu,\nu) \geq \int L_j d\mu + \int \psi d\nu.$$

To obtain the opposite inequality, just use (39) and (40), and observe that, if the mass of μ is $T-t$ and the mass of ν is 1, then, for each (χ,ϕ), we have

$$\int -\chi - L d\mu \leq (T-t) \sup(-\chi - L),$$

as well as

$$\int \psi - \phi d\nu \leq \sup(\psi - \phi)\,,$$

so that

$$\begin{aligned} & -\int \chi d\mu - \int \phi d\nu - h(\chi,\phi) \\ \leq & \int L d\mu + \int \psi d\nu + (T-t)\sup(-\chi - L) + \sup(\psi-\phi) - h(\chi,\phi) \\ = & \int L d\mu + \int \psi d\nu. \end{aligned}$$

This completes the proof of (41).
To compute the Legendre transform of g, observe that if $(\mu,\nu) \notin \mathcal{M}_2$ then there exist functions $(\hat{\chi},\hat{\phi}) \in \mathcal{C}$ such that

$$\int \hat{\chi} d(\mu_* - \mu) - \int \hat{\phi} d(\nu_* - \nu) \neq 0\,, \tag{44}$$

thanks to the fact that $(\mu_*,\nu_*) \in \mathcal{M}_2$. Possibly changing the signs of $\hat{\chi}$ and $\hat{\psi}$, we may suppose that the quantity in (44) is positive. Then, take $\hat{\chi}_j := j\hat{\chi}$ and $\hat{\phi}_j := j\hat{\phi}$. We get that

$$\lim_{j\to+\infty} \int \hat{\chi}_j d(\mu_* - \mu) - \int \hat{\phi}_j d(\nu_* - \nu) \ = \ +\infty\,.$$

The above considerations show that

$$g^*(\mu,\nu) = \inf_{\chi,\phi} - \int \chi d(\mu_* - \mu) + \int \phi d(\nu_* - \nu) = \begin{cases} 0 & \text{if } (\mu,\nu) \in \mathcal{M}_2 \\ -\infty & \text{otherwise,} \end{cases}$$

which gives (42). ☐

We now complete the proof of Proposition 6 by arguing as follows. We note that the function g (resp., h), as defined above, is concave (resp., convex) and upper (resp., lower) semicontinuous (we remark indeed that $\mathcal{C}$ is a closed convex set).
Then, by the Legendre-Fenchel-Rockafellar Theorem (see, e.g., [Vil03] or [Gom06]),

$$\sup(g-h) \,=\, \inf(h^* - g^*)\,.$$

Consequently, (29), (36) and Lemma 7 yield:

$$\begin{aligned}
\tilde{u}^\epsilon &= \inf_{(\mu,\nu)\in\mathcal{M}_1\cap\mathcal{M}_2} \int L\,d\mu + \int \psi\,d\nu_t^\epsilon \\
&= \inf_{(\mu,\nu)} h^* - g^* \\
&= \sup_{(\chi,\phi)} g - h \\
&= \sup_{(\chi,\phi)\in\mathcal{C}} (t-T) \sup_{(x',X,v,s)} (-\chi - L) - \sup_{(x',X)} (\psi-\phi) + \int \phi\,d\nu_t^\epsilon \\
&= \sup_{\phi\in C^2_\Lambda} (t-T) \sup_{(x',X,v,s)} (-\tilde{A}\phi - \partial_t\phi - L) - \sup_{(x',X)} (\psi-\phi) + \int \phi\,d\nu_t^\epsilon \\
&= \sup_{\phi\in C^2_\Lambda} (t-T) \sup_{(x',X,s)} (\tilde{H} - \partial_t\phi) - \sup_{(x',X)} (\psi-\phi) + \int \phi\,d\nu_t^\epsilon\,,
\end{aligned}$$

which ends the proof of of Proposition 6. ☐

The result to come shows that (38) is optimized when $X = \frac{x'}{\epsilon}$:

Proposition 8. *Suppose that ν_t^ϵ is supported in the set $X = \frac{x'}{\epsilon}$. Then*

$$\begin{aligned}
&\tilde{u}^\epsilon(x,t,\omega) \\
&= \sup_{\phi(x',X,s)\in C^2_\Lambda(\mathbb{R}^n\times\mathbb{R}^n\times[t,T])} (T-t) \left[\inf_{(x',X=\frac{x'}{\epsilon},s)\in\mathbb{R}^n\times\mathbb{R}^n\times(t,T)} \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right) \right] \\
&\quad + \inf_{x'} \left[\psi(x') - \phi(x', \frac{x'}{\epsilon}, T) \right] + \\
&\quad + \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x',X,t)d\nu_t^\epsilon \\
&= \sup_{\varphi(x',s)\in C^2_\Lambda(\mathbb{R}^n\times[t,T])} (T-t) \Bigg[\inf_{(x',s)\in\mathbb{R}^n\times(t,T)} \partial_t\varphi \\
&\quad - H\left(D^2_x\varphi, D_x\varphi, x', \frac{x'}{\epsilon}, \epsilon, \omega \right) \Bigg] \\
&\quad + \inf_{x'} [\psi(x') - \varphi(x',T)] \\
&\quad + \int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi(x',t)d\nu_t^\epsilon
\end{aligned}$$

Proof. Let $\varkappa$ be as in the proof of Proposition 3, satisfying the additional uniformity hypothesis: for any $\delta > 0$ there exists $\kappa(\delta) > 0$ such that

$$\inf_{|y|\geq\delta} \varkappa(y) \geq \kappa(\delta).$$

Then

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \varkappa\left(X - \frac{x'}{\epsilon}\right) d\nu_t^\epsilon = 0. \tag{45}$$

By contradiction, assume that the first equality claimed in Proposition 8 is false, hence, by (38), that

$$\begin{aligned}\tilde{u}^\epsilon(x,t,\omega) \ < \ & \sup_{\phi(x',X,s)} \inf_{(x',X=\frac{x'}{\epsilon},s)\in\mathbb{R}^n\times\mathbb{R}^n\times(t,T)} \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right) \\ & + \inf_{x',X}\left[\psi(x') - \phi(x', \frac{x'}{\epsilon}, T)\right] \\ & + \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x', X, t) d\nu_t^\epsilon\,,\end{aligned}$$

for some $x \in \mathbb{R}^n$.

This would mean that for some $\delta > 0$, possibly depending on x, that

$$\begin{aligned} & \tilde{u}^\epsilon(x,t,\omega) \\ = \ & \sup_{\phi(x',X,s)} \inf_{(x',X,s)\in\mathbb{R}^n\times\mathbb{R}^n\times(t,T),\ |X-\frac{x'}{\epsilon}|>\delta} \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right) \\ & + \inf_{x',X,\ |X-\frac{x'}{\epsilon}|>\delta} \left[\psi(x') - \phi(x', X, T)\right] + \\ & + \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(x', X, t) d\nu_t^\epsilon.\end{aligned} \tag{46}$$

We replace $\phi(x', X, t)$ by $\phi(x', X, t) + \lambda(t - T)\varkappa\left(X - \frac{x'}{\epsilon}\right)$ in (46), with $\lambda > 0$. Since, due to the observation next to (37), this does not change $\tilde{H}$, we have that the first infimum in (46) appears as

$$\begin{aligned}\lambda\varkappa\left(X - \frac{x'}{\epsilon}\right) &+ \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right) \\ &\geq \lambda\kappa(\delta) + \partial_t\phi - \tilde{H}\left(D^2\phi, D\phi, x', X, \epsilon, \omega\right),\end{aligned}$$

and the other infimum, as well as the integral are unchanged, due to (45). By sending $\lambda \to \infty$ we would have a contradiction, and thus that the infimum in (38) must be achieved when $X = \frac{x'}{\epsilon}$.

This shows the first equality in Proposition 8.

The second one follows by taking $\varphi(x', t) := \phi(x', \frac{x'}{\epsilon}, t)$ and by observing that ν_t^ϵ is supported on $\{X = \frac{x'}{\epsilon}\}$, thanks to (45), so that

$$\int \phi(x', X, t)\, d\nu_t^\epsilon = \int \varphi(x', t)\, d\nu_t^\epsilon\,.$$

□

From this point on, we will assume Assumption 6. With this, we are now in the position to improve Proposition 4:

Proposition 9. *$\tilde{u}^\epsilon = u^\epsilon$ in $\mathbb{R}^n \times [t,T] \times \Omega$.*

Proof. By (22), ν_t^ϵ is supported in $X = \frac{x'}{\epsilon}$. Thus, by Proposition 8 and (9), we have that

$$\begin{aligned}
\tilde{u}^\epsilon &\geq (T-t) \inf_{x',s} \partial_t v_\delta^\epsilon - H\left(D_x^2 v_\delta^\epsilon, D_x v_\delta^\epsilon, x', \frac{x'}{\epsilon}, \epsilon, \omega\right) \\
&\quad + \inf_{x'} \left[\psi(x') - v_\delta^\epsilon(x',T)\right] + v_\delta^\epsilon(x,t) \\
&\geq -\delta + 0 + v_\delta^\epsilon(x,t) .
\end{aligned}$$

Sending $\delta \to 0$, and using (10), we conclude that $\tilde{u}^\epsilon \geq u^\epsilon$. The converse inequality is assured by Proposition 4. □

Finally, we present a result that converts the constrained minimization problem of (29) and (30) into an unconstrained problem. The idea, quite standard in linear programming, consists in using the approximate solution v_δ^ϵ of the dual problem as a Lagrange multiplier so to remove the constraints.

Proposition 10 (Unconstrained minimization)**.** *Fix $\delta > 0$, and let v_δ^ϵ be as in Assumption 6. Then,*

$$\begin{aligned}
&\left| u^\epsilon(x,t,\omega) - \left[\inf_\mu \int_{\mathbb{R}^n\times U\times[t,T]} L(x', \frac{x'}{\epsilon}, v, \omega) + A v_\delta^\epsilon + \partial_t v_\delta^\epsilon \, d\mu + \int v_\delta^\epsilon d\nu_t^\epsilon\right]\right| \\
&\leq \delta(T-t) + \sup_{s\in[t,T]} \left|u^\epsilon(x,s,\omega) - v_\delta^\epsilon(x,s,\omega)\right| , \qquad (47)
\end{aligned}$$

where the above infimum is taken over all measures $\mu(x',v,s)$ on $\mathbb{R}^n \times U \times [t,T]$, with total mass $T-t$.

Proof. By (6),

$$L(x', \frac{x'}{\epsilon}, v, \omega) + A v_\delta^\epsilon \geq -H(D_x^2 v_\delta^\epsilon, D_x v_\delta^\epsilon, x', \frac{x'}{\epsilon}, \epsilon, \omega),$$

and so, integrating with respect to any measure μ with total mass $T-t$,

$$\int L(x', \frac{x'}{\epsilon}, v, \omega) + A v_\delta^\epsilon + \partial_t v_\delta^\epsilon \, d\mu \geq \int \partial_t v_\delta^\epsilon - H(D_x^2 v_\delta^\epsilon, D_x v_\delta^\epsilon, x', \frac{x'}{\epsilon}, \epsilon, \omega) \, d\mu \geq -\delta\,(T-t) ,$$

thanks to (9). This and the definition of ν_t^ϵ, equation (22), imply that

$$\begin{aligned}
&u^\epsilon(x,t,\omega) - \left[\inf_\mu \int_{\mathbb{R}^n\times U\times[t,T]} L(x', \frac{x'}{\epsilon}, v, \omega) + A v_\delta^\epsilon + \partial_t v_\delta^\epsilon \, d\mu + \int v_\delta^\epsilon d\nu_t^\epsilon\right] \\
&\leq \delta(T-t) + u^\epsilon(x,t,\omega) - v_\delta^\epsilon(x,t,\omega) . \qquad (48)
\end{aligned}$$

On the other hand, let $\kappa > 0$, and suppose that $(\mathbf{x}(s), \mathbf{v}(s))$ is an almost optimal trajectory for (5), for $s \in [t, T]$, so that

$$u^\epsilon \geq \iint L(x', \frac{x'}{\epsilon}, v, \omega) d\mu_o^\epsilon(x', v, s; \omega) + E\psi\left(\mathbf{x}(T)\right) - \kappa\,, \tag{49}$$

where the measure μ_o^ϵ on $\mathbb{R}^n \times U \times [t, T]$ is defined by

$$\int \phi(x', v, s)\, d\mu_o^\epsilon := E \int_t^T \phi(\mathbf{x}(s), \mathbf{v}(s), s)\, ds\,, \tag{50}$$

for any $\phi \in C_*(\mathbb{R}^n \times U \times [t, T])$.
Notice that $\mu_o^\epsilon(\mathbb{R}^n \times U \times [t, T]) = T - t$, therefore

$$\begin{aligned} &\inf_\mu \int L(x', \frac{x'}{\epsilon}, v, \omega) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu \\ \leq & \int L(x', \frac{x'}{\epsilon}, v, \omega) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu_o^\epsilon. \end{aligned}$$

From this, (49) and (22), we get that

$$\begin{aligned} & \kappa + u^\epsilon(x, t, \omega) - \left[\inf_\mu \int_{\mathbb{R}^n \times U \times [t,T]} L(x', \frac{x'}{\epsilon}, v, \omega) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu + \int v_\delta^\epsilon d\nu_t^\epsilon\right] \\ \geq\ & E\psi\left(\mathbf{x}(T), \frac{\mathbf{x}(T)}{\epsilon}, T\right) - \int Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu_o^\epsilon - \int v_\delta^\epsilon\, d\nu_t^\epsilon \\ =\ & E\psi\left(\mathbf{x}(T), \frac{\mathbf{x}(T)}{\epsilon}, T\right) - \int Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu_o^\epsilon - v_\delta^\epsilon(x, t, \omega)\,. \end{aligned} \tag{51}$$

Since, by (50) and Dynkin's Formula,

$$\int Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu_o^\epsilon + v_\delta^\epsilon(x, t, \omega) = Ev_\delta^\epsilon(\mathbf{x}(T), T),$$

we deduce from (51) that

$$\begin{aligned} & \kappa + u^\epsilon(x, t, \omega) - \left[\inf_\mu \int_{\mathbb{R}^n \times U \times [t,T]} L(x', \frac{x'}{\epsilon}, v, \omega) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu + \int v_\delta^\epsilon d\nu_t^\epsilon\right] \\ \geq\ & E\left[\psi\left(\mathbf{x}(T)\right) - v_\delta^\epsilon(\mathbf{x}(T), T, \omega)\right] = E\left[u^\epsilon(\mathbf{x}(T), T, \omega) - v_\delta^\epsilon(\mathbf{x}(T), T)\ \right]. \end{aligned}$$

From this estimates, sending $\kappa \to 0$, and (48), we obtain (47). $\square$

We remark that Proposition 10 and Assumption 6 imply that

$$\begin{aligned} & \lim_{\delta \to 0} \inf_\mu \int_{\mathbb{R}^n \times U \times [t,T]} L\left(x', \frac{x'}{\epsilon}, v, \omega\right) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon\, d\mu + \int v_\delta^\epsilon d\nu_t^\epsilon \\ =\ & u^\epsilon(x, t, \omega) \end{aligned} \tag{52}$$

uniformly in $(x, t, \omega) \in \mathbb{R}^n \times [0, T] \times \Omega$, where the the infimum is taken over all measures $\mu(x', v, s)$ on $\mathbb{R}^n \times U \times [t, T]$, with total mass $T - t$.

7. Stationarity and Effective Lagrangian

From this section on we will assume Assumption 7. We observe that, as a consequence of Assumption 7 and equation (6), the Hamiltonian is also stationary, that is

$$H(M,p,x,X+Y,\epsilon,\omega)=H(M,p,x,X,\epsilon,\tau_Y\omega)\,. \tag{53}$$

Under reasonable hypothesis one can show that the minimizing measures μ^ϵ are tight in the variable x, however, the same is false for the variable X. Therefore, we need a procedure to compactify the space in such a way that we can use weak convergence tools. The stationary setting provides a convenient way to achieve this, and for that we introduce the following construction: given a measure $\mu^\epsilon(x',X,v,s;\omega)$ in $\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]$, we construct a measure $\hat\mu^\epsilon(x',v,s,\omega)$ in $\mathbb{R}^n\times U\times[t,T]\times\Omega$ defined in the following way:

$$\int \phi(x',v,s,\omega)d\hat\mu^\epsilon \;:=\; \iint \phi(x',v,s,\tau_X\omega)d\mu^\epsilon d\omega\,, \tag{54}$$

for any bounded measurable function ϕ in $\mathbb{R}^n\times U\times[t,T]\times\Omega$.
Conversely, given a measure $\hat\mu^\epsilon(x',v,s,\omega)$, one can build a measure $\bar{\hat\mu}^\epsilon(x',X,v,s,\omega)$ in $\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]\times\Omega$ in the following way:

$$\int \phi(x',X,v,s,\omega)d\bar{\hat\mu}^\epsilon \;:=\; \int \phi(x',0,v,s,\omega)d\hat\mu^\epsilon\,, \tag{55}$$

for any bounded measurable ϕ in $\mathbb{R}^n\times\mathbb{R}^n\times U\times[t,T]\times\Omega$.
Of course, with definitions similar to (54) and (55) (by omitting the dependence on v) one can construct from ν_T^ϵ measures $\hat\nu_T^\epsilon$ and $\bar{\hat\nu}_T^\epsilon$.
Note that if ϕ is stationary, i.e., if

$$\phi(x',X,v,s,\tau_Y\omega)\;=\;\phi(x',X+Y,v,s,\omega)\,,$$

then

$$\int\phi\,d\bar{\hat\mu}^\epsilon=\int\phi\,d\mu^\epsilon d\omega\,.$$

Proposition 11. *(i) Let μ^ϵ and ν_T^ϵ be a sequence of minimizing measures for* (29)*, corresponding to the initial trace ν_t^ϵ, according to* (30)*. Then, the quantity in* (29) *is uniformly bounded.*
(ii) Let μ^ϵ and ν_T^ϵ be a sequence of minimizing measures satisfying (30) *for which the quantity in* (29) *is uniformly bounded. Let $\hat\mu^\epsilon$ be as defined in* (54)*, and define similarly $\hat\nu_T^\epsilon$. Then, there exist subnets, which, by abuse of notation are still denoted as $\hat\mu^\epsilon$ and $\hat\nu_T^\epsilon$, and measures $\rho(x',s)$ and $\theta(v,\omega;x',s)$ such that*

$$\hat\mu^\epsilon(x',v,s,\omega)\rightharpoonup\theta(v,\omega;x',s)\,\rho(x',s)\,. \tag{56}$$

Furthermore, for each fixed (x',s), θ is a probability measure in $U\times\Omega$. Additionally, we have measures $\bar\nu_T$ and $\bar\nu_t$ on $\mathbb{R}^n$, defined in the following way:

$$\int\varphi(x')\,d\bar\nu_T(x')\;:=\;\lim_{\epsilon\to0}\int\varphi(x')\,d\hat\nu_T^\epsilon(x',X,\omega) \tag{57}$$

and

$$\int \varphi(x')\,d\bar{\nu}_t(x') := \lim_{\epsilon\to 0}\int \varphi(x')\,d\hat{\nu}_t^{\epsilon}(x',X,\omega)\,, \tag{58}$$

for[2] *any* $\varphi \in C_c(\mathbb{R}^n)$.

Proof. The claim in (i) easily follows by comparing with the control $\mathbf{v} := 0$.
We now prove the claim in (ii). Consider the control $\mathbf{v} := 0$ and the corresponding trajectory $\mathbf{x}_0(t;\omega)$. By comparing with the measures defined in (23) and (24), and by exploiting (4), we get that there exists a constant C such that

$$\begin{aligned} C &\geq \int\left[\int_t^T L\left(\mathbf{x}_0(s;\omega),\frac{\mathbf{x}_0(s;\omega)}{\epsilon},0,\omega\right)\,ds + 2\sup|\psi|\right]d\omega \\ &\geq \int L(x,0,v,\omega)\,d\hat{\mu}^{\epsilon} \\ &\geq \left[(T-t)-\hat{\mu}^{\epsilon}\Big(\mathbb{R}^n\times(U\setminus B_R)\times[t,T]\Big)\right]\inf L \\ &\qquad +\mu^{\epsilon}\Big(\mathbb{R}^n\times(U\setminus B_R)\times[t,T]\Big)\inf_{|v|\geq R} L. \end{aligned}$$

From the growth conditions in v of L (see equation (4) in Assumption 1) we thus conclude that, as $R\to\infty$,

$$\hat{\mu}^{\epsilon}\Big(\mathbb{R}^n\times(U\setminus B_R)\times[t,T]\Big)\to 0.$$

Furthermore, we know, from Proposition 5 that

$$\hat{\mu}^{\epsilon}\Big((\mathbb{R}^n\setminus B_R)\times U\times[t,T]\Big)\to 0,$$

as $R\to\infty$, which then implies that

$$\hat{\mu}^{\epsilon}\Big((\mathbb{R}^n\setminus B_R)\times(U\setminus B_R)\times[t,T]\Big)\to 0,$$

as $R\to\infty$, and then implies tightness. Therefore, the desired claim follows from Assumptions 8 and 3.
From Proposition 5, it is also clear that the measure in $\mathbb{R}^n\times\Omega$ defined by

$$E\int_{\mathbb{R}^n}\varphi(x',\omega)d\nu_T^{\epsilon}$$

is tight, and therefore we have (57). The weak convergence in (58) is simply a consequence of (22), as then $\bar{\nu}_t = \delta_x(x')$. □

From this point on, we will also assume that the scaling hypotheses in Assumption 9 hold, that is $\sigma = \epsilon^{\alpha}\sigma_0$, with $\alpha = 1, \frac{1}{2}$ or 0 (resp. cases A, B or C). In each of these cases, we define the generalized Mather problem in the following way: given $\varpi\in\mathbb{R}^n$

[2] A standard observation is that the moment estimates in Proposition 5 imply that (57) and (58) also hold for any continuous $\varphi(x')$ growing at infinity less than $|x'|^p$ once they hold for compactly supported ones.

and $\Xi \in \mathrm{Mat}(n \times n)$ we look, for each fixed x, at probability measures $\eta(v, \omega; x, \varpi, \Xi)$ on $U \times \Omega$, which minimize

$$\overline{L}(x, \varpi, \Xi) := \int L(x, 0, v, \omega) d\eta(v, \omega; x, \varpi, \Xi) \tag{59}$$

under the constraints, which depend on the scaling hypothesis put forward in Assumption 9, given by:

A. In this first case, the Effective Lagrangian $\overline{L}(x, \varpi)$ will not depend on Ξ, and we take as constraints the rotation vector

$$\int f(x, 0, v, \omega) d\eta = \varpi, \tag{60}$$

and the holonomy constraint

$$\int f(x, 0, v, \omega) \cdot D_X \phi(0, \omega) d\eta = 0, \tag{61}$$

for all stationary functions $\phi(X, \omega)$, which are C^1 in the first variable. Note that in this scaling the contribution of the diffusion vanishes.

B. In this intermediate scaling, the Effective Lagrangian will also not depend on Ξ, and we require, similarly, the rotation vector

$$\int f(x, 0, v, \omega) d\eta = \varpi, \tag{62}$$

and the holonomy constraint, which in this case has a contribution from the diffusion

$$\int f(x, 0, v, \omega) \cdot D_X \phi(0, \omega) + \frac{\sigma_0 \sigma_0^T(x, 0, v, \omega)}{2} : D_X^2 \phi(0, \omega) d\eta = 0, \tag{63}$$

for all C^2 (in the first variable) stationary functions $\phi(X, \omega)$.

C. In this scaling, in which diffusion dominates, we require both an average diffusion coefficient Ξ, a rotation vector ϖ

$$\begin{aligned} \frac{1}{2} \int \sigma_0 \sigma_0^T d\eta = \Xi, \\ \int f(x, 0, v, \omega) d\eta = \varpi, \end{aligned} \tag{64}$$

and the holonomy constraint, which does not have a drift term since the diffusion dominates

$$\int \frac{\sigma_0 \sigma_0^T(x, 0, v, \omega)}{2} : D_X^2 \phi(0, \omega) d\eta = 0, \tag{65}$$

for all C^2 in X stationary functions $\phi(X, \omega)$.

Given the limiting measures ρ and θ constructed in Proposition 11, we define, for each $(x', s) \in \mathbb{R}^n \times [t, T]$,

$$\bar{\varpi}(x', s) := \int f(x', 0, v, \omega) \, d\theta(\omega, v; x', s) \, . \tag{66}$$

Also, in case C, we define

$$\bar{\Xi}(x', s) := \frac{1}{2} \int \sigma_0 \sigma_0^T(x', 0, v, \omega)\, d\theta(\omega, v; x', s). \tag{67}$$

It turns out that $\rho(x', s)$ almost everywhere, the measure $\theta(v, \omega; x', s)$ constructed in Proposition 11 satisfies the holonomy constraints corresponding to cases A-C. More precisely:

Proposition 12. *For almost any $(x', s) \in \mathbb{R}^n \times [t, T]$ with respect to the measure $\rho(x', s)$, the measure $\theta(v, \omega; x', s)$ satisfies conditions (61), (63) or (65), and conditions (60), (62) or (64) with ϖ (resp., Ξ) replaced by $\bar{\varpi}(x', s)$ (resp., $\bar{\Xi}(x', s)$), according to cases A-C. Moreover, in cases A and B, if we define the measure $\varrho(x', \varpi, s)$ on $\mathbb{R}^n \times \mathbb{R}^n \times [t, T]$ by $\varrho(x', \varpi, s) := \delta_{\bar{\varpi}(x',s)}(\varpi)\rho(x', s)$, we have that*

$$\int \varpi \cdot D_x\varphi(x', s) + \partial_t\varphi(x', s) d\varrho = \int \varphi(x', T) d\bar{\nu}_T - \int \varphi(x', t) d\bar{\nu}_t, \tag{68}$$

for all C^1 functions φ.
Analogously, in case C, if we define the measure $\varrho(x', \varpi, \Xi, s)$ on $\mathbb{R}^n \times \mathbb{R}^n \times \mathrm{Mat}(n \times n) \times [t, T]$ by $\varrho(x', \varpi, \Xi, s) := \delta_{\bar{\varpi}(x',s)}(\varpi)\delta_{\bar{\Xi}(x',s)}(\Xi)\rho(x', s)$, we have that

$$\begin{aligned} &\int \varpi \cdot D_x\varphi(x', s) + \Xi : D_x^2\varphi(x', s) + \partial_t\varphi(x', s) d\varrho \\ &= \int \varphi(x', T) d\bar{\nu}_T - \int \varphi(x', t) d\bar{\nu}_t, \end{aligned} \tag{69}$$

for all C^2 functions φ.

Proof. As usual, given a, $b \in \mathbb{R}^n$, we define $a \otimes b \in \mathrm{Mat}(n \times n)$ by $(a \otimes b)_{ij} := a_i b_j$. Let $\phi(X, \omega) : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be C^2 in X and stationary, let $\alpha(x', s) \in C(\mathbb{R}^n \times [t, T])$ and consider the function

$$(x', X, s, \omega) \mapsto \alpha(x', s)\, \phi(X, \omega)\,.$$

By (21), (12) and (13), we see that

$$\begin{aligned} &\int \left[\frac{1}{\epsilon} f(x', 0, v, \omega) \cdot D_X\phi(0, \omega) + \frac{1}{2\epsilon^2}(\sigma\sigma^T)(x', 0, \epsilon, v, \omega) : D_X^2\phi(0, \omega)\right] \alpha(x', s)\, d\hat{\mu}^\epsilon \\ &\quad + \int \left[\partial_t\alpha + f(x', 0, v, \omega) \cdot D_x\alpha + \frac{1}{2}(\sigma\sigma^T)(x', 0, \epsilon, v, \omega) : D_x^2\alpha\right] \phi(0, \omega)\, d\hat{\mu}^\epsilon \\ &\quad + \frac{1}{\epsilon} \int (\sigma\sigma^T)(x', 0, \epsilon, v, \omega) : (D_x\alpha \otimes D_X\phi(0, \omega))\, d\hat{\mu}^\epsilon \\ &= \iint \alpha\phi\, d\nu_T^\epsilon\, d\omega - \iint \alpha\phi\, d\nu_t^\epsilon\, d\omega\,. \end{aligned} \tag{70}$$

We first consider cases A and B. We now take $a := 1$ in case A or $a := 1/2$ in case B and we multiply (70) by ϵ, thus obtaining

$$\begin{aligned}
&\int \left[f(x',0,v,\omega)\cdot D_X\phi(0,\omega) + \frac{\epsilon^{2a-1}}{2}(\sigma_0\sigma_0^T)(x'0,v,\omega) : D_X^2\phi(0,\omega)\right] \alpha(x',s)\,d\hat{\mu}^\epsilon \\
&\qquad + \epsilon \int \left[\partial_t\alpha + f(x',0,v,\omega)\cdot D_x\alpha + \frac{1}{2}(\sigma\sigma^T)(x',0,\epsilon,v,\omega) : D_x^2\alpha \right] \phi(0,\omega)\,d\hat{\mu}^\epsilon \\
&\qquad + \epsilon^{2a} \int (\sigma_0\sigma_0^T)(x'0,v,\omega) : (D_x\alpha \otimes D_X\phi(0,\omega))\,d\hat{\mu}^\epsilon \\
&= \epsilon \left[\iint \alpha\phi\,d\nu_T^\epsilon\,d\omega - \iint \alpha\phi\,d\nu_t^\epsilon\,d\omega \right].
\end{aligned}$$

By sending $\epsilon \to 0$ via Proposition 11, it follows that

$$\begin{aligned}
\iint \Big[f(x',0,v,\omega)\cdot D_X\phi(0,\omega) + \frac{c}{2}(\sigma_0\sigma_0^T)(x'0,v,\omega) : D_X^2\phi(0,\omega) \Big] \\
d\theta(v,\omega;x',s)\,\alpha(x',s)\,d\rho(x',s) \;=\; 0\,,
\end{aligned}$$

where $c = 0$ in case A, and $c = 1$ in case B.
Since α is arbitrary, we conclude that

$$\int \left[f(x',0,v,\omega)\cdot D_X\phi(0,\omega) + \frac{c}{2}(\sigma_0\sigma_0^T)(x'0,v,\omega) : D_X^2\phi(0,\omega)\right] d\theta(v,\omega;x',s) \;=\; 0$$

for almost any (x',s) with respect to ρ, thence conditions (61) or (63) are fulfilled, according to cases A and B.
Multiplying (70) by ϵ^2 and arguing as above, one shows that (65) holds in case C.
The fact that θ satisfies (60), (62) or (64) with ϖ replaced by $\bar{\varpi}(x',s)$ follows at once by (66) and (67).
To prove (68) and (69), it is sufficient to plug $\phi := 1$ in (70) and send $\epsilon \to 0$. □

Proposition 13. *In the notation of Proposition 12, we have that*

$$\int \overline{L}\,d\varrho \le \iint L\,d\theta\,d\rho\,. \tag{71}$$

Proof. We use the notation of case C, since cases A and B are analogous and easier, to compute that

$$\begin{aligned}
&\int \overline{L}(x',\varpi,\Xi)\,d\varrho \\
&= \iint L(x',0,v,\omega)\,d\eta(v,\omega;x',\varpi,\Xi)\,d\varrho(x',\varpi,\Xi,s) \\
&= \iint L(x',0,v,\omega)\,d\eta(v,\omega;x',\bar{\varpi}(x',s),\bar{\Xi}(x',s))\,d\rho(x',s) \\
&\le \iint L(x',0,v,\omega)\,d\theta(v,\omega;x',s)\,d\rho(x',s)\,,
\end{aligned}$$

as we wished. □

We remark that, for some choices of ϖ and Ξ, the constrained minimization problem given in (59) may be ill posed, since there might be no measure fulfilling conditions (61), (63) or (65), and conditions (60), (62) or (64), according to cases A-C. However, one of the consequences of Proposition 12 is that such a problem is always well defined for some particular choice of ϖ and Ξ, i.e., at least for $\varpi := \bar{\varpi}(x', s)$ and $\Xi := \bar{\Xi}(x', s)$. In case, for some (x^*, ϖ^*, Ξ^*), there are no measures fulfilling conditions (61), (63) or (65), and conditions (60), (62) or (64), according to cases A-C, we define $\overline{L}(x^*, \varpi^*, \Xi^*)$ to be $+\infty$.

Proposition 14. *The Effective Lagrangian $\overline{L}$ is lower semicontinuous.*

Note that this result, via Yosida regularization (see, e.g., Theorem 2.64 in [Att84]), gives that $\overline{L}$ may be approximated monotonically from below by continuous functions.

Proof (of Proposition 14). Let $(x'_n, \varpi_n, \Xi_n) \to (x', \varpi, \Xi)$ and consider the corresponding optimal measures η_n on $U \times \Omega$.
Let

$$\overline{L}_n := \overline{L}(x'_n, \varpi_n, \Xi_n) = \int L(x'_n, 0, v, \omega) d\eta_n(v, \omega; x'_n, \varpi_n, \Xi_n)\,.$$

Without loss of generality, we may suppose that $\overline{L}_n$ is bounded by above, say $\overline{L}_n \leq K$ (otherwise, either $\liminf \overline{L}_n = \infty$, in which case there is nothing to show, or $\liminf \overline{L}_n < \infty$ and by extracting a subsequence we are in the previous case). But then η_n is tight by (4) and so the result follows from Assumption 8. $\square$

We now observe some growth conditions of $\overline{L}$:

Proposition 15. *In cases A and B, we have that $|\varpi|^2 \leq C(1 + \overline{L}(x', \varpi))$.*
In case C, we have that $|\varpi|^2 + |\Xi|^p \leq C(1 + \overline{L}(x', \varpi, \Xi))$.

Proof. We will make use of Assumptions 4 and 15. By (60), (62) or (64) (according to the cases),

$$\begin{aligned} |\varpi| \leq \int |f(x', 0, v, \omega)| d\eta \;&\leq\; 1 \cdot \left(\int |f(x', 0, v, \omega)|^2 d\eta \right)^{1/2} \\ &\leq\; C \left(\int 1 + L(x', 0, v, \omega) d\eta \right)^{1/2} \\ &=\; C(1 + \overline{L})^{1/2}\,. \end{aligned}$$

In case C, an analogous argument applied to (64) gives

$$\begin{aligned} 2|\Xi| \leq \int |\sigma_0 \sigma_0^T(x', 0, v, \omega)| d\eta \;&\leq\; 1 \cdot \left(\int |\sigma_0 \sigma_0^T(x', 0, v, \omega)|^p d\eta \right)^{1/p} \\ &\leq\; C \left(\int 1 + L(x', 0, v, \omega) d\eta \right)^{1/p} \\ &=\; C(1 + \overline{L})^{1/p}\,, \end{aligned}$$

as desired. $\square$

8. Homogenized problems

In this section, we discuss the homogenized problem, which is built using the Effective Lagrangian constructed in the previous section.

Let ν_t be as in (22) and let $\overline{L}(x', \varpi, \Xi)$ or $\overline{L}(x', \varpi)$ be defined as in the previous section. We consider the following problems:

A or B. We look for measures $\varrho(x', \varpi, s)$ on $\mathbb{R}^n \times \mathbb{R}^n \times [t, T]$ and $\bar{\nu}_T(x')$ on $\mathbb{R}^n$ which minimize

$$\bar{u}(x,t) := \int \overline{L}(x', \varpi) d\varrho + \int \psi d\bar{\nu}_T, \tag{72}$$

under the constraint in (68).

C. We look for measures $\varrho(x', \varpi, \Xi, s)$ on $\mathbb{R}^n \times \mathbb{R}^n \times \mathrm{Mat}(n \times n) \times [t, T]$ and $\bar{\nu}_T(x')$ on $\mathbb{R}^n$ which minimize

$$\bar{u}(x,t) := \int \overline{L}(x', \varpi, \Xi) d\varrho + \int \psi d\bar{\nu}_T, \tag{73}$$

under the constraint in (69).

This variational problem is well posed, according to the result to come:

Proposition 16. *There exist minimizing measures $\varrho(x', \varpi, s)$ on $\mathbb{R}^n \times \mathbb{R}^n \times [t, T]$ and $\bar{\nu}_T(x')$ on $\mathbb{R}^n$, as requested above. Moreover,*

$$|\bar{u}(x,t)| \le C(1 + \|\psi\|_\infty + T - t) \tag{74}$$

for any $x \in \mathbb{R}^n$ and $t \le T$

Proof. The idea of the proof is that the measures built in Propositions 11 and 12 give an upper bound on the action and a natural candidate for the minimization process (though maybe not the minimum itself). The minimization process then converges due to the moment estimates of Proposition 5 and the lower semicontinuity of $\overline{L}$ given by Proposition 14.

Following are the details of the proof. First, we consider the process

$$d\mathbf{x}^\epsilon = f\Big(\mathbf{x}^\epsilon, \frac{\mathbf{x}^\epsilon}{\epsilon}, 0, \omega\Big) dt + \sigma\Big(\mathbf{x}^\epsilon, \frac{\mathbf{x}^\epsilon}{\epsilon}, \epsilon, 0, \omega\Big) dW_t$$

with $\mathbf{x}^\epsilon(t) = x$ and let μ_ϵ and ν_ϵ be as in (23) and (24) with $v(s) := 0$. Then, by Assumption 14,

$$\begin{aligned} & \int L d\mu_\epsilon + \int \psi d\nu_\epsilon \\ = \; & E \int_t^T L\Big(\mathbf{x}^\epsilon(s), \frac{\mathbf{x}^\epsilon(s)}{\epsilon}, 0, \omega\Big) ds + \psi\Big(\mathbf{x}^\epsilon, \frac{\mathbf{x}^\epsilon}{\epsilon}\Big) \le C(1 + \|\psi\|_\infty + T - t) \end{aligned} \tag{75}$$

and so the second claim in Proposition 11 holds true. Consequently, the claim in Proposition 12 holds true. In particular, the measure μ_ϵ induces in the limit a measure ϱ satisfying (68) and (69). By (71) and (75), we get that

$$\int \overline{L} d\varrho + \int \psi d\bar{\nu}_T \le C(1 + \|\psi\|_\infty + T - t) \tag{76}$$

that is, the quantities in (72) and (73) are not always infinite.
This allow us to use the direct method in the calculus of variations, that is, we take a minimizing sequence of measures, say ϱ_n and $\bar\nu_{T,n}$. We apply to these measures the moment estimates of Proposition 5 (with respect to the notation in Proposition 5, we replace what there is called μ^ϵ, ν_T^ϵ, f, $\sigma\sigma^T$ and L with the quantities denoted here by ϱ_n, $\bar\nu_{T,n}$, ϖ, Ξ and $\overline{L}$). Note that the measures ϱ_n and $\bar\nu_{T,n}$ satisfy a constraint of the type in (30), thanks to (68) or (69), that the uniform bound on the action of the type in (32) is assured here by (76) and that the growth condition in Assumption 4 is given here by Proposition 15.
Then, the tightness given by Proposition 5 and Assumption 8 gives the existence of the desired minimal measures, since, by Proposition 14, $\overline{L}$ is lower semicontinuous.
Also, the uniform bound in (76) easily implies (74). □

Proposition 17. *The measure ϱ, as introduced in Proposition 12, satisfies* (68) *in cases A and B, and* (69) *in case C.*

Proof. Let $\varphi(x',t)$ be a smooth function. We use (21), (12) and (13) to gather that

$$\begin{aligned}&\int \partial_t\varphi(x',s) + f(x',0,v,\omega)\cdot D_x\varphi(x',s) + \frac{\sigma\sigma^T(x',0,\epsilon,v,\omega)}{2} : D_x^2\varphi(x',s)\, d\hat\mu^\epsilon\\ =\ &\iint \varphi(x',T)\, d\nu_T^\epsilon\, d\omega - \iint \varphi(x',t)\, d\nu_t^\epsilon\, d\omega\,.\end{aligned}$$

Therefore, we use Proposition 11 to send $\epsilon\to 0$, and we obtain

$$\begin{aligned}&\iint \partial_t\varphi(x',s) + f(x',0,v,\omega)\cdot D_x\varphi(x',s)\\ &+\frac{\sigma'\sigma'^T(x',0,v,\omega)}{2} : D_x^2\varphi(x',s)\, d\theta(v,\omega;x',s)\, d\rho(x',s)\\ =\ &\int \varphi(x',T)\, d\bar\nu_T^\epsilon - \int \varphi(x',t)\, d\bar\nu_t^\epsilon\,,\end{aligned}\tag{77}$$

where $\sigma' := 0$ in cases A and B, and $\sigma' := \sigma_0(x',0,v,\omega)$ in case C.
By means of Proposition 12, we know that

$$\begin{aligned}&\frac12\int \sigma_0\sigma_0^T\, d\theta(v,\omega;x's) = \bar\Xi(x',s)\\ \text{and}\qquad &\int f(x,0,v,\omega)\, d\theta(v,\omega;x's) = \bar\varpi(x',s)\,.\end{aligned}$$

Accordingly, the left hand side of (77) is equal to

$$\int \partial_t\varphi(x',s) + \bar\varpi(x',s)\cdot D_x\varphi(x',s) + c\bar\Xi(x',s) : D_x^2\varphi(x',s)\, d\rho(x',s)\,,$$

where $c := 0$ in cases A and B, and $c := 1$ in case C.
The latter quantity equals

$$\int \partial_t\varphi + \varpi\cdot D_x\varphi(x',s) + c\Xi(x',s) : D_x^2\varphi(x',s)\, d\varrho\,,$$

thus yielding the desired result. ☐

9. Uniform estimates

This section is dedicated to proving certain uniform estimates that allow us to obtain convergence of u^ϵ to a non-random limit. For that we will assume, in addition to the previous ones, Assumptions 11-13. We will prove two types of estimates: the first type concerns a-priori estimates for the modulus of continuity of u^ϵ which are uniform in ϵ – these estimates apply fairly standard viscosity solutions and optimal control techniques and are used to establish that the solutions u^ϵ converge to a limit pointwise in ω. The second type of estimates show that the dependence on ω decreases as $\epsilon \to 0$, and therefore establish that any limit of u^ϵ is non-random.

Proposition 18. *We have*

$$\|u_t\|_\infty \leq \limsup_{\epsilon\to 0} \sup_x \left| H\left(D^2\psi, D\psi, x, \frac{x}{\epsilon}, \epsilon, \omega\right)\right| . \tag{78}$$

Proof. By Assumption 10, we may assume that both the terminal data ψ and the solution u^ϵ are smooth. Define $w(x,t) = u_t(x, T-t)$. Then, w satisfies the parabolic equation

$$w_t + Lw = 0, \tag{79}$$

where L is the elliptic operator

$$Lw = D_M H : D^2 w + D_p H \cdot Dw,$$

and $w(x,0) = H(D^2\psi, D\psi, x, \frac{x}{\epsilon})$. Note that the fact that L is elliptic (thence (79) is parabolic) is a consequence of (14).
Since equation (79) has maximum principle (see, for instance, Proposition 4.1 and Theorem 4.2 in [DiB95] or page 165 in [Shu92]), we obtain the required estimate. ☐

Remark 19. We observe that the right hand side of (78) is finite, due to (8).

Proposition 20. $\|Du^\epsilon\|_\infty$ *is uniformly bounded in* ϵ.

Proof. We will prove separately the cases in which H satisfies P1 or P2 of Assumption 11.
Case 1. If H satisfies P1 then we have, from Proposition 18

$$H(Du^\epsilon, x, X) \leq C,$$

which by (16) implies that $\|Du^\epsilon\|_\infty$ is bounded.
Case 2. By Assumption 10, we may suppose that u^ϵ and ψ are smooth. Consider the following rescaled version of the solution

$$U(\tau, y) = \frac{1}{\epsilon} u^\epsilon(\epsilon\tau, \epsilon y).$$

Then

$$\|U\|_\infty < +\infty \tag{80}$$

because of Proposition 18 and Remark 19, and U satisfies the equation

$$-U_\tau + H\Big(\frac{1}{\epsilon}D^2U, DU, \epsilon y, y, \epsilon, \omega\Big) = 0,$$

with terminal data $\frac{1}{\epsilon}\psi(\epsilon y)$ at $\tau = \frac{T}{\epsilon}$. A simple argument shows that u^ϵ is Lipschitz if and only if so is U, with the same Lipschitz bound. According to our smoothness hypothesis, we must now show that DU is bounded.
Let $w = |DU|^2/2$. A simple computation yields

(81)
$$\begin{aligned} w_\tau &= DU \cdot DU_\tau \\ &= \frac{1}{\epsilon}D_iUD_{M_{jk}}HD^3_{ijk}U + D_iUD_{p_j}HD^2_{ij}U + \epsilon D_iUD_{x_i}H + D_iUD_{X_i}H \\ &= \frac{1}{\epsilon}D_{M_{jk}}HD^2_{jk}w - \frac{1}{\epsilon}D_{M_{jk}}HD^2_{ij}UD^2_{ik}U + D_{p_j}HD_jw + \epsilon D_iUD_{x_i}H + D_iUD_{X_i}H. \end{aligned}$$

If w admits an interior maximum, we have $w_t = 0$ and $Dw = 0$ there, and

$$D_{M_{jk}}HD^2_{jk}w \geq 0\,,$$

due to (14). Thus,

$$-\frac{1}{\epsilon}D_{M_{jk}}HD^2_{ij}UD^2_{ik}U + \epsilon D_iUD_{x_i}H + D_iUD_{X_i}H \leq 0.$$

But then P2 of Assumption 11 and (80) imply that DU is uniformly bounded in this case.
We now modify the above argument in case w does not attain its maximum. For this, we observe that $\|w\|_\infty < +\infty$, thanks to (15). Hence, if, for a small q (even smaller than ϵ), we define $w^{(q)}(y) := w(y) - q\ln(1+|y|^2)$ we have that $w^{(q)}$ has an interior maximum, say at y_q. Then, $w^{(q)}_t(y_q) = 0$, $Dw^{(q)}_t(y_q) = 0$ and $D_{M_{jk}}HD^2_{jk}w^{(q)}_t(y_q) \geq 0$, by (14). As a result, $w_t(y_q) = 0$, $|Dw_t(y_q)| \leq Cq$ and $D_{M_{jk}}HD^2_{jk}w^{(q)}_t(y_q) \geq -Cq$, for some $C > 0$. Substituting in (81), we obtain that

$$-\frac{1}{\epsilon}D_{M_{jk}}HD^2_{ij}UD^2_{ik}U + \epsilon D_iUD_{x_i}H + D_iUD_{X_i}H \leq C_\epsilon q$$

at the point y_q, for some $C_\epsilon > 0$, due to (8).
Then, P2 of Assumption 11 and (80) yield that

$$1 \geq C_\epsilon q \geq F(DU(y_q), U_\tau(y_q)) \geq \inf_{|\zeta|\leq\|U_\tau\|_\infty} F(DU(y_q), \zeta)\,,$$

as long as q is small enough.
From this and (17), we infer that $|DU(y_q)|$ is uniformly bounded for small q. But then, $w \leq w(y_q)$ is uniformly bounded, and so is DU, as desired. $\square$

As a consequence, we have

Proposition 21. *Fixed any $\omega \in \Omega$, through an appropriate subsequence, we have that $u^\epsilon(x,t,\omega)$ converges to $u^\sharp(x,t,\omega)$ locally uniformly in $(x,t) \in \mathbb{R}^n \times [0,T]$, for some $u^\sharp : \mathbb{R}^n \times [0,T] \to \mathbb{R}$. Also,*

$$\lim_{\epsilon\to 0} \int_\Omega u^\epsilon(x,t,\omega)\, d\omega \;=\; u^0(x,t)\,,$$

for some $u^0 : \mathbb{R}^n \times [0,T] \to \mathbb{R}$.

Proof. This follows from the Ascoli-Arzelà Theorem, since $\int_\Omega u^\epsilon(x,t,\omega)\, d\omega$ is equicontinuous, thanks to Propositions 18 and 20, and equibounded, due to the terminal condition. □

Proposition 22. *We have that*

$$|u^\epsilon(x,t,\tau_y\omega) - u^\epsilon(x,t,\omega)| \le C\epsilon|y|\,. \tag{82}$$

Moreover,

$$|\int u^\epsilon(x-y,t,\omega)\, d\omega - \int u^\epsilon(x,t,\omega)\, d\omega| \le |y| \Big[\|D\psi\|_\infty + C(T-t) \Big]\,. \tag{83}$$

Proof. Notice that

$$-u^\epsilon_t(x,t,\tau_y\omega) + H\Big(D^2u^\epsilon(x,t,\tau_y\omega), Du^\epsilon(x,t,\tau_y\omega), x, \frac{x}{\epsilon}, \epsilon, \tau_y\omega\Big) = 0.$$

By (53),

$$-u^\epsilon_t(x,t,\tau_y\omega) + H(D^2u^\epsilon(x,t,\tau_y\omega), Du^\epsilon(x,t,\tau_y\omega), x, y + \frac{x}{\epsilon}, \epsilon, \omega) = 0.$$

Set $v(x,y,t;\omega) = u^\epsilon(x-\epsilon y, t, \tau_y\omega)$. Then, v solves

$$-v_t + H(D^2v, Dv, x-\epsilon y, \frac{x}{\epsilon}, \epsilon, \omega) = 0.$$

As before, by Assumption 10, we may assume that v is smooth. Then set

$$w = \frac{\partial v}{\partial y_k}\,.$$

Thus,

$$-w_t + D_M H : D^2 w + D_p H \cdot Dw - \epsilon D_{x_k} H = 0.$$

By Assumption 12, the ellipticity condition in (14) and the maximum principle (see, e.g., Proposition 4.1 and Theorem 4.2 in [DiB95]), we conclude that

$$\|w(x,t)\|_\infty \le \epsilon\|D\psi\|_\infty + \epsilon C(T-t),$$

which implies

$$|u^\epsilon(x-\epsilon y, t, \tau_y\omega) - u^\epsilon(x,t,\omega)| \le |y| \left[\epsilon\|D\psi\|_\infty + \epsilon C(T-t)\right]\,. \tag{84}$$

Furthermore, the uniform Lipschitz continuity of u^ϵ (recall Proposition 20) implies that

$$|u^\epsilon(x-\epsilon y, t, \tau_y\omega) - u^\epsilon(x,t,\tau_y\omega)| \le C\epsilon|y|.$$

The latter two estimates yield the claim in (82).
Then, (83) follows from (84) by replacing ϵy with y and recalling that τ_y is measure preserving, according to Assumption 7. □

Proposition 23. *We have*

$$\limsup_{\epsilon\to 0} \sup_{\substack{x\in\mathbb{R}^n \\ t\in[0,T] \\ \omega_0,\omega_1\in\Omega}} |u^\epsilon(x,t,\omega_0)-u^\epsilon(x,t,\omega_1)| \to 0.$$

Proof. To prove the above estimate, observe that

$$\begin{aligned}|u^\epsilon(x,t,\omega_0)-u^\epsilon(x,t,\omega_1)| \le {} & |u^\epsilon(x,t,\omega_0)-u^\epsilon(x,t,\tau_y\omega_0)| \\ & + |u^\epsilon(x,t,\tau_y\omega_0)-u^\epsilon(x,t,\omega_1)|.\end{aligned}$$

Let $\delta_1>0$ be fixed, and suppose that for every ϵ there exists $y=y(x,t,\epsilon,\delta_1)$ such that $|y|\le M$, with M depending only on δ_1 (and so independent of ϵ), for which

$$\sup_{\substack{x\in\mathbb{R}^n \\ t\in[0,T] \\ \omega_0,\omega_1\in\Omega}} |u^\epsilon(x,t,\tau_y\omega_0)-u^\epsilon(x,t,\omega_1)| \le \frac{\delta_1}{2}. \tag{85}$$

Then, from Proposition 22, for ϵ sufficiently small, we have

$$\sup_{\substack{x\in\mathbb{R}^n \\ t\in[0,T] \\ \omega_0\Omega \\ |y|\le M}} |u^\epsilon(x,t,\omega_0)-u^\epsilon(x,t,\tau_y\omega_0)| \le \frac{\delta_1}{2},$$

which then, by sending $\delta_1\to 0^+$ yields the estimate. Therefore our task now consists in establishing (85).
To this extent, fix $\delta>0$ and choose y as in Assumption 13. Then

$$\tilde u^\epsilon(x,t) := u^\epsilon(x,t,\omega_1)+\delta(T-t)$$

satisfies

$$\begin{aligned}0 &= -u^\epsilon_t(x,t,\omega_1)+H\Big(D^2u^\epsilon(x,t,\omega_1),D\tilde u^\epsilon(x,t,\omega_1),x,\frac{x}{\epsilon},\epsilon,\omega_1\Big)\\ &= -\delta-\tilde u^\epsilon_t+H\Big(D^2\tilde u,D\tilde u,x,\frac{x}{\epsilon},\epsilon,\omega_1\Big)\\ &\le -\tilde u^\epsilon_t+H\Big(D^2\tilde u,D\tilde u,x,\frac{x}{\epsilon},\epsilon,\tau_y\omega_0\Big),\end{aligned}$$

that is

$$-\tilde u^\epsilon_t+H\Big(D^2\tilde u,D\tilde u,x,\frac{x}{\epsilon},\epsilon,\tau_y\omega_0\Big)\ge 0\,.$$

Then, $u^\epsilon-\tilde u^\epsilon$ is a subsolution of a linear parabolic equation, due to (14), which then implies, by the comparison principle for viscosity solutions (see Proposition 4.1 and Theorem 4.2 in [DiB95]) that

$$u^\epsilon(x,t,\tau_y\omega_0)\le \tilde u^\epsilon(x,t).$$

Accordingly, we obtain (85) by possibly exchanging the roles of $\tau_y\omega_0$ and ω_1, as desired. □

By collecting the results in this section, we get the following result:

Proposition 24. *Through an appropriate subsequence,*

$$\lim_{\epsilon\to 0} u^\epsilon(x,t,\omega) \,=\, u^0(x,t)\,,$$

uniformly in $K\times[0,T]\times\Omega$, for any compact set $K\subset\mathbb{R}^n$.

Proof. Fix $\omega_0\in\Omega$. Then, by Proposition 21, $u^{\epsilon_j}(x,t,\omega_0)$ converges to $u^\sharp(x,t,\omega_0)$ locally uniformly in (x,t). Let $u^0(x,t):=u^\sharp(x,t,\omega_0)$. Then,

$$\begin{aligned}
&\lim_{j\to+\infty}\sup_{(x,t,\omega)\in K\times[0,T]\times\Omega}|u^{\epsilon_j}(x,t,\omega)-u^0(x,t)|\\
\le\;&\lim_{j\to+\infty}\sup_{(x,t,\omega)\in K\times[0,T]\times\Omega}|u^{\epsilon_j}(x,t,\omega)-u^{\epsilon_j}(x,t,\omega_0)|\\
+\;&\lim_{j\to+\infty}\sup_{(x,t)\in K\times[0,T]}|u^{\epsilon_j}(x,t,\omega_0)-u^\sharp(x,t,\omega_0)|\\
\le\;&0+0\,,
\end{aligned}$$

where Proposition 23 has been exploited. □

10. Convergence

In this section, we establish that the limit of the viscosity solutions u^ϵ agrees with $\bar u$, which was defined in (72). We divide the proof into two parts, a lower bound and an upper bound.

Proposition 25 (Lower Bound)**.** *$u^0\ge\bar u$.*

Proof. Consider the measures ϱ and θ constructed in section 7. Let also μ^ϵ and ν_T^ϵ be the measures optimizing (5).
Then, we use Proposition 9, (29) (11), (54), Proposition 11, (71), Proposition 17, (72) and (73) to obtain that

$$\begin{aligned}
u^0(x,t)&=\lim_{\epsilon\to0}\int_\Omega u^\epsilon(x,t,\omega)\,d\omega\\
&=\lim_{\epsilon\to0}\left[\iint L(x',X,v,s,\omega)d\mu^\epsilon d\omega+\iint\psi(x')d\nu_T^\epsilon\,d\omega\right]\\
&=\left[\lim_{\epsilon\to0}\int L(x',0,v,s,\omega)d\hat\mu^\epsilon\right]+\int\psi d\bar\nu_T\\
&=\iint L(x',0,v,s,\omega)d\theta\,d\rho+\int\psi d\bar\nu_T\\
&\ge\int\overline{L}(x',\varpi,\Xi)d\varrho+\int\psi d\bar\nu_T=\bar u(x,t).
\end{aligned}$$

□

In order to obtain the opposite inequality to Proposition 25, we need the result in Proposition 24.

Proposition 26 (Upper bound)**.** *$u^0=\bar u$.*

Proof. Let $\delta > 0$ and v_δ^ϵ be as in Assumption 6. We denote by o_δ quantities (possibly depending on (x,t,ω)) which tend to zero as $\delta \to 0$.
Let $\kappa > 0$. We consider almost optimal measures $(\varrho_\kappa(x', \varpi, \Xi, s), \bar\nu_{T,\kappa}(x'))$ for (72) or (73), and an almost optimal family of probability measure $\eta_\kappa(v,\omega; x', s, \varpi, \Xi)$ for (59), that is, we suppose that $\varrho_\kappa(x', \varpi, \Xi, s)$ almost everywhere

$$\overline{L}(x', \varpi, \Xi) \geq \int L(x, 0, v, \omega) d\eta_\kappa(v, \omega; x, \varpi, \Xi) - \kappa,$$

and that, in case A or B,

$$\bar u(x,t) \geq \int \overline{L}(x', \varpi) d\varrho_\kappa + \int \psi d\bar\nu_{T,\kappa} - \kappa,$$

whereas, in case C,

$$\bar u(x,t) \geq \int \overline{L}(x', \varpi, \Xi) d\varrho_\kappa + \int \psi d\bar\nu_{T,\kappa} - \kappa.$$

We construct a trial measure $\tilde\mu_\kappa(x', v, s, \omega)$ on $\mathbb{R}^n \times U \times [t,T] \times \Omega$ defined in the following way:

$$\int_{\mathbb{R}^n \times U \times [t,T] \times \Omega} \phi(x', v, s, \omega) d\tilde\mu_\kappa := \iint \phi(x', v, s, \tau_{-\frac{x'}{\epsilon}}\omega) d\eta_\kappa d\varrho_\kappa\,, \tag{86}$$

for any $\phi \in C_c(\mathbb{R}^n \times U \times [t,T] \times \Omega)$.
Note that, by taking $\varphi(s) := (s - T)$ in (68) or (69), we obtain that

$$\int d\varrho_\kappa = T - t\,,$$

and so, since η_κ is a probability measure, it follows that $\tilde\mu_\kappa$ has total mass $T-t$.
In particular, the total mass of $\tilde\mu_\kappa$ is finite and so, recalling Assumption 3, we can slice $\tilde\mu_\kappa$ as

$$\tilde\mu_\kappa(x', v, s, \omega) =: \zeta_\kappa(x', v, s; \omega)\, \Theta_\kappa(\omega)\,, \tag{87}$$

where $\Theta_\kappa(\omega)$ is a probability measure, and ζ_κ is, for Θ_κ-almost any $\omega \in \Omega$ fixed, a measure of finite total mass. Let $m(\omega)$ be the total mass of ζ_κ for a fixed ω. By construction,

$$1 = \frac{1}{T-t}\int d\tilde\mu_\kappa = \frac{1}{T-t}\iint d\zeta_\kappa\, d\Theta_\kappa = \int \frac{m(\omega)}{T-t} d\Theta_\kappa(\omega)\,. \tag{88}$$

Let also

$$\hat\zeta_\kappa := \frac{T-t}{m(\omega)}\zeta_\kappa\,,$$

and note that $\hat\zeta$ has total mass $T-t$ for any fixed ω.
Then, by Proposition 10,

$$o_\delta + u^\epsilon \quad \leq \quad \int L(x', \frac{x'}{\epsilon}, v, \omega) + Av_\delta^\epsilon + \partial_t v_\delta^\epsilon d\hat\zeta_\kappa + \int v_\delta^\epsilon d\nu_t^\epsilon\,,$$

and so, by (87) and (88),

$$\begin{aligned} o_\delta &+ \int \frac{m(\omega)}{T-t} u^\epsilon(x,t,\omega) d\Theta_\kappa(\omega) \\ &\le \int L(x', \frac{x'}{\epsilon}, v, \omega) d\tilde{\mu}_\kappa + \int A v_\delta^\epsilon + \partial_t v_\delta^\epsilon d\tilde{\mu} + \iint \frac{m(\omega)}{T-t} v_\delta^\epsilon d\nu_t^\epsilon d\Theta_\kappa. \end{aligned}$$

Making use of (11) and (59), we get that

$$\begin{aligned} \int L(x', \frac{x'}{\epsilon}, v, \omega) d\tilde{\mu}_\kappa &= \int L(x', \frac{x'}{\epsilon}, v, \tau_{-\frac{x'}{\epsilon}}\omega) d\eta_\kappa d\varrho_\kappa \\ &= \iint L(x', 0, v, \omega) d\eta_\kappa d\varrho_\kappa \\ &\le \int \overline{L}(x', \varpi, \Xi) d\varrho_\kappa + C\kappa, . \end{aligned}$$

for some suitable constant $C > 0$. Similarly, by (19), (86), (12), (13), (60), (62) and (64),

$$\begin{aligned} \int A\varphi \, d\tilde{\mu}_\kappa &= \int f(x', \frac{x'}{\epsilon}, v, \omega) \cdot D_x\varphi + \frac{\sigma\sigma^T}{2}(x', \frac{x'}{\epsilon}, v, \omega) : D_x^2\varphi \, d\tilde{\mu}_\kappa \\ &= \iint f(x', \frac{x'}{\epsilon}, v, \tau_{-\frac{x'}{\epsilon}}\omega) \cdot D_x\varphi \\ &\quad + \frac{\sigma\sigma^T}{2}(x', \frac{x'}{\epsilon}, v, \tau_{-\frac{x'}{\epsilon}}\omega) : D_x^2\varphi \, d\eta_\kappa \, d\varrho_\kappa \\ &= \iint f(x', 0, v, \omega) \cdot D_x\varphi + \frac{\sigma\sigma^T}{2}(x', 0, v, \omega) : D_x^2\varphi \, d\eta_\kappa \, d\varrho_\kappa \\ &= \int \varpi \cdot D_x\varphi + \Xi^\epsilon : D^2\varphi \, d\varrho_\kappa \,, \end{aligned}$$

for any $\varphi(x') \in C^2(\mathbb{R}^n)$, with

$$\Xi^\epsilon(x') := \frac{1}{2} \int \sigma\sigma^T(x', 0, v, \omega) \, d\eta_\kappa \,.$$

Thus, exploiting (68) and (69),

$$\begin{aligned} \int A v_\delta^\epsilon + \partial_t v_\delta^\epsilon d\tilde{\mu}_\kappa &= \int \varpi \cdot D_x v_\delta^\epsilon + \Xi^\epsilon : D_x^2 v_\delta^\epsilon + \partial_t v_\delta^\epsilon d\varrho_\kappa \\ &= \int v_\delta^\epsilon d\bar{\nu}_T - \int v_\delta^\epsilon d\bar{\nu}_t + \int (\Xi^\epsilon - \Xi^*) : D_x^2 v_\delta^\epsilon d\varrho_\kappa \,, \end{aligned}$$

with $\Xi^* := 0$ in cases A and B and $\Xi^* = \Xi = \Xi^\epsilon$ in case C.
Note that

$$\lim_{\epsilon \to 0} \Xi^\epsilon - \Xi^* = 0 \tag{89}$$

uniformly in (x, v, ω) in all the cases A-C.

Also, by (72) and (73),

$$\bar{u}(x,t) \geq \int \overline{L}(x', \varpi, \Xi) d\varrho_\kappa + \int \psi d\bar{\nu}_{T,\kappa} - \kappa \,.$$

By collecting the above estimates, we conclude that

$$\begin{aligned} o_\delta + \int \frac{m(\omega)}{T-t} u^\epsilon(x,t,\omega) d\Theta(\omega) \\ \leq \bar{u}(x,t) + \int (v_\delta^\epsilon - \psi) d\bar{\nu}_T + \iint \frac{m(\omega)}{T-t} v_\delta^\epsilon d\nu_t^\epsilon \, d\Theta - \int v_\delta^\epsilon d\bar{\nu}_t \\ + \int (\Xi^\epsilon - \Xi^*) : D_x^2 v_\delta^\epsilon d\varrho_\kappa + C\kappa \,. \end{aligned}$$

That is, by Assumption 6,

$$\begin{aligned} & o_\delta + \int \frac{m(\omega)}{T-t} u^\epsilon \, d\Theta \\ \leq \ & \bar{u} + \int (u^\epsilon - \bar{\psi}) \, d\bar{\nu}_T + \iint \frac{m(\omega)}{T-t} u^\epsilon d\nu_t^\epsilon \, d\Theta \\ & - \int u^\epsilon d\bar{\nu}_t + \int (\Xi^\epsilon - \Xi^*) : D_x^2 v_\delta^\epsilon d\varrho + C\kappa \,. \end{aligned}$$

We now send $\epsilon \to 0$ and we obtain

$$o_\delta + u^0 \leq \bar{u} + C\kappa \,,$$

thanks to Proposition 24 and formulas (58), (88) and (89).
We now send δ and κ to zero, getting $u^0 \leq \bar{u}$. This and Proposition 25 yield the desired claim. □

11. Improved Moment Estimates

In this section we improve the moment estimates of section 5 in order to show that for small time "optimal trajectories don't go too far". These estimates will be essential in establishing that the homogenized limit is a viscosity solution of a suitable equation. Here, we will take a small time step $h \in (0, T)$. We start with an auxiliary result:

Lemma 27. *Consider the diffusion*

$$d\mathbf{x}^0 = f\Big(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega\Big)dt + \sigma\Big(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega\Big)dW_t,$$

with initial condition $\mathbf{x}^0(T-h) = x$. *Then, we have*

$$|E[\mathbf{x}^0(T-h) - \mathbf{x}^0(T)]| \leq Ch,$$

and

$$E|\mathbf{x}^0(T-h) - \mathbf{x}^0(T)|^2 \leq Ch,$$

where the constant C *can be chosen independent of* ϵ, h *and* x.

Proof. Fix $h > 0$, let $T - h \le t \le T$. By Dynkin's Formula we have

$$\frac{d}{dt}E[\mathbf{x}^0(T-h) - \mathbf{x}^0(t)] = -Ef\Big(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega\Big), \tag{90}$$

which, by Assumption 14, yields the first part of the claim.
Similarly, we have

$$\begin{aligned}\frac{d}{dt}E|\mathbf{x}^0(T-h) - \mathbf{x}^0(t)|^2 =& 2E\left[(\mathbf{x}^0(T-h) - \mathbf{x}^0(t))f(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega)\right] \\ &+ \operatorname{tr} E[\sigma(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega)\sigma^T(\mathbf{x}^0, \frac{\mathbf{x}^0}{\epsilon}, 0, \omega)]. \end{aligned} \tag{91}$$

Thus, if we set $A(t) = E|\mathbf{x}^0(T-h) - \mathbf{x}^0(t)|^2$ we have

$$\frac{d}{dt}A \le A + C, \qquad A(T-h) = 0.$$

A Gronwall estimate yields $|A(t)| \le Ch$. $\square$

Lemma 28. *Consider an optimal diffusion $\mathbf{x}(t)$, for $T - h \le t \le T$, with initial condition $\mathbf{x}(T-h) = x$. Let*

$$M := E\int_{T-h}^{T} L\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)dt.$$

Then, we have

$$|E[\mathbf{x}(T-h) - \mathbf{x}(T)]| \le M^{1/2}Ch^{1/2} + Ch,$$

and

$$E|\mathbf{x}(T-h) - \mathbf{x}(T)|^2 \le MCh + Ch,$$

where C can be chosen independently of ϵ, h and x.

Proof. Using again (90) and Assumption 15 we obtain

$$\begin{aligned} E[\mathbf{x}(T-h) - \mathbf{x}(T)] &= E\int_{T-h}^{T} f\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)dt \\ &\le h^{1/2}\left(E\int_{T-h}^{T}\Big|f\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)\Big|^2 dt\right)^{1/2} \\ &\le Ch^{1/2}\left(E\int_{T-h}^{T} L\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big) + Cdt\right)^{1/2} \\ &\le CM^{1/2}h^{1/2} + Ch. \end{aligned}$$

Similarly, by (91), we have

$$\begin{aligned} \frac{d}{dt}E|\mathbf{x}(T-h) - \mathbf{x}(t)|^2 =& 2E\Big[(\mathbf{x}(T-h) - \mathbf{x}(t))f\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)\Big] \\ &+ \operatorname{tr} E\Big[\sigma\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)\sigma^T\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big)\Big] \\ \le& Ch^{-1}E|\mathbf{x}(T-h) - \mathbf{x}(t)|^2 + ChEL\Big(\mathbf{x}, \frac{\mathbf{x}}{\epsilon}, \mathbf{v}, \omega\Big) + C. \end{aligned}$$

Thus, if we set $A(t)=E|\mathbf{x}(T-h)-\mathbf{x}(t)|^2$ we have, by a Gronwall estimate, that $|A(t)|\leq CMh+Ch$. ☐

Now we use the two previous results to prove that the integral of the Lagrangian along an optimal trajectory is not too big:

Proposition 29. *Consider an optimal diffusion* $\mathbf{x}$ *for* $T-h\leq t\leq T$*, minimizing*

$$E\int_{T-h}^{T} L\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt+\psi(\mathbf{x}(T)),$$

with initial condition $\mathbf{x}(T-h)=x$*. Then there exists a constant* C *independent of* h*,* ϵ *and* x *such that*

$$E\int_{T-h}^{T} L\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt\leq Ch.$$

Proof. Consider the diffusion

$$d\mathbf{x}^0=f\Big(\mathbf{x}^0,\frac{\mathbf{x}^0}{\epsilon},0,\omega\Big)dt+\sigma\Big(\mathbf{x}^0,\frac{\mathbf{x}^0}{\epsilon},0,\omega\Big)dW_t,$$

with $\mathbf{x}(T-h)=x$. We have, by the optimality of $\mathbf{x}$, that

$$\begin{aligned}&E\left[\int_{T-h}^{T} L\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt+\psi(\mathbf{x}(T))\right]\\ &\qquad\leq E\left[\int_{T-h}^{T} L\Big(\mathbf{x}^0,\frac{\mathbf{x}^0}{\epsilon},0,\omega\Big)dt+\psi(\mathbf{x}^0(T))\right].\end{aligned}$$

Moreover, the fact that $\mathbf{x}(T-h)=\mathbf{x}^0(T-h)=x$, via a second order Taylor expansion of ψ, gives that

$$\begin{aligned}&\psi(\mathbf{x}^0(T))-\psi(\mathbf{x}(T))\\ \leq\;& D\psi(x)\cdot(\mathbf{x}^0(T)-\mathbf{x}^0(T-h))-D\psi(x)\cdot(\mathbf{x}(T)-\mathbf{x}(T-h))\\ &+C\Big(|\mathbf{x}^0(T)-\mathbf{x}^0(T-h)|^2+|\mathbf{x}(T)-\mathbf{x}(T-h)|^2\Big).\end{aligned}$$

By the previous estimates and Lemmata 27 and 28,

$$\begin{aligned}M=E\left[\int_{T-h}^{T} L\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt\right]\leq&E\left[\int_{T-h}^{T} L\Big(\mathbf{x}^0,\frac{\mathbf{x}^0}{\epsilon},0,\omega\Big)dt+\psi(\mathbf{x}^0(T))-\psi(\mathbf{x}(T))\right]\\ \leq&\, C(1+M)h+CM^{1/2}h^{1/2},\end{aligned}$$

which implies the desired bound. ☐

We will also need the following estimate:

Proposition 30. *Fix* $r>0$*. Consider the optimal diffusion*

$$d\mathbf{x}=f\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt+\sigma\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dW_t,$$

with $\mathbf{x}(T-h)=x$*. Then, as* $h\to 0^+$*, we have*

$$P(|\mathbf{x}(T)-x|>r)=o(h).$$

Proof. Write
$$\mathbf{x}=\mathbf{y}+\mathbf{z},$$
where
$$d\mathbf{y}=f\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dt,\qquad d\mathbf{z}=\sigma\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)dW_t,$$
and $\mathbf{y}(T-h)=x$, $\mathbf{z}(T-h)=0$.
We have
$$\begin{aligned}\frac{d}{dt}E|\mathbf{y}-x|^2&=2E(\mathbf{y}-x)f\Big(\mathbf{x},\frac{\mathbf{x}}{\epsilon},\mathbf{v},\omega\Big)\\&\le 2\left(E|\mathbf{y}-x|^2\right)^{1/2}\left(E|f|^2\right)^{1/2},\end{aligned}$$
which implies
$$\frac{d}{dt}\left(E|\mathbf{y}-x|^2\right)^{1/2}\le C\left(E|f|^2\right)^{1/2}.$$
Thus, using again Assumption 15,
$$\begin{aligned}\left(E|\mathbf{y}(t)-x|^2\right)^{1/2}&\le C\int_{T-h}^{T}\left(E|f|^2\right)^{1/2}dt\\&\le Ch^{1/2}\left(\int_{T-h}^{T}E|f|^2dt\right)^{1/2}\\&\le Ch^{1/2}\left(E\int_{T-h}^{T}L+Cdt\right)^{1/2}\le Ch,\end{aligned}$$
due to Proposition 29, and so
$$E|\mathbf{y}(t)-x|^2\le Ch^2.$$
Accordingly, by Chebychev's inequality, we have
$$P\Big(|\mathbf{y}(T)-x|>\frac{r}{2}\Big)\le C\frac{h^2}{r^2}. \tag{92}$$
To handle the second term, we use Itô's formula (see, e.g., page 105 in [Eva06b]) and we get
$$\begin{aligned}&\frac{d}{dt}Ee^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}\\&=-E\frac{\lambda|z|^2e^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}}{(t+2h-T)^2}+\frac{1}{2}E\sigma\sigma^T:\left[\frac{2\lambda I}{t+2h-T}e^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}+\frac{\lambda^2z\otimes z}{(t+2h-T)^2}e^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}\right]\\&\le E\sigma\sigma^T:\left[\frac{\lambda I}{t+2h-T}e^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}\right]\\&\le\frac{C}{h}Ee^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}},\end{aligned}$$
if λ is sufficiently small and $t\in[T-h,T]$.
Thus, recalling that $\mathbf{z}(T-h)=0$, we obtain by a Gronwall estimate that
$$Ee^{\frac{\lambda|\mathbf{z}|^2}{t+2h-T}}\le C,$$

for all $T-h \leq t \leq T$. Again by Chebychev's inequality,

$$P\Big(|\mathbf{z}(T)| \geq \frac{r}{2}\Big) \leq Ce^{-\frac{\lambda r^2}{2h}}, \tag{93}$$

which is exponentially small as $h \to 0$. The desired result then follows from (92) and (93). ☐

We will need a final moment estimate for the optimal measures of the homogenized problem. Such moment estimate will only be in use in the proof of Proposition 33 and it requires the following additional condition:

Assumption 16. We assume that for any fixed $(x_0, \varpi_0, \sigma_0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathrm{Mat}(n \times n)$ there exists a map $x \mapsto \mathcal{S}_{x_0,\varpi_0,\sigma_0}(x) = (\mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(x), \mathcal{S}^{\sigma}_{x_0,\varpi_0,\sigma_0}(x)) \in \mathbb{R}^n \times \mathrm{Mat}(n \times n)$ with the properties listed below.
Let $\Xi_0 = \frac{1}{2}\sigma_0\sigma_0^T$, and $\mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(x) = \frac{1}{2}\mathcal{S}^{\sigma}_{x_0,\varpi_0,\sigma_0}(x)\left(\mathcal{S}^{\sigma}_{x_0,\varpi_0,\sigma_0}(x)\right)^T$. Then

1. $\mathcal{S}_{x_0,\varpi_0,\sigma_0}(x_0) = (\varpi_0, \sigma_0)$;
2. the mapping $x \mapsto \mathcal{S}_{x_0,\varpi_0,\sigma_0}(x)$ is globally bounded and globally Lipschitz;
3. the mapping $x \mapsto \overline{L}(\mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(x), \mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(x), x)$ is globally Lipschitz.

◀

Assumption 16 is obviously fulfilled, by choosing $\mathcal{S}_{x_0,\varpi_0,\sigma_0}(x)$ to be identically equal to (ϖ_0, σ_0), when $\overline{L}$ is globally Lipschitz in x and, in particular, if $\overline{L}$ does not depend on x at all – this is the case, for instance, if L does not depend on x (though it may depend on $X = x/\epsilon$).

Proposition 31. *Suppose that $T \in (0,1]$ is conveniently small. Let ϱ and ν_T be optimal measures for* (72) *under the constraint in* (68) *in cases A and B, or optimal measures for* (73) *under the constraint in* (69) *in case C, with $t = 0$.*
Let also Assumption 16 hold. Then,

$$\left|\int (x'-x)d\nu_T\right| \leq CT, \tag{94}$$

$$\int |x'-x|^2 d\varrho \leq CT^2 \tag{95}$$

and

$$\int |x'-x|^2 d\nu_T \leq CT. \tag{96}$$

Furthermore,

$$\left|\int \overline{L} d\varrho\right| \leq CT. \tag{97}$$

Proof. First of all, we observe that,

$$|\varpi|^2 + |\Xi| + |\Xi|^p \leq C(1 + \overline{L}(x', \varpi, \Xi)) \tag{98}$$

thanks to Proposition 15.

Also, by taking $\varphi(x', s) = x'_j - x_j$ in the holonomy constraint (68) or (69) (according to the case), with $j = 1, \ldots, n$, we obtain the identity

$$\int \varpi d\varrho = \int x' - x d\nu_T. \tag{99}$$

By taking $\varphi = \frac{1}{2}(T-s)|x'-x|^2$, we obtain

$$\int (T-s)(\varpi \cdot (x'-x) + \operatorname{tr} \Xi) - \frac{1}{2}|x'-x|^2 d\varrho = 0. \tag{100}$$

Similarly, take $\varphi = \frac{1}{2}|x'-x|^2$ to get

$$\int \varpi \cdot (x'-x) + \operatorname{tr} \Xi \, d\varrho = \int |x'-x|^2 d\nu_T. \tag{101}$$

By taking $\varphi = 1$ and $\varphi = T - s$ we also see that the total masses of ν_T and ϱ are 1 and T, respectively. Thus, using (99), the Hölder Inequality and (98), we conclude that

$$\left| \int (x'-x) d\nu_T \right| \le C T^{1/2} \left(\int |\varpi|^2 d\varrho \right)^{1/2} \le CT + CT^{1/2} \left(\int \overline{L} d\varrho \right)^{1/2}. \tag{102}$$

Also, from (100) and a scaled Hölder Inequality,

$$\int |x'-x|^2 d\varrho \le CT \int |\varpi||x'-x| + |\Xi| d\varrho \le \frac{1}{2} \int |x'-x|^2 d\varrho + CT \int |\varpi|^2 + |\Xi| d\varrho$$

and so, by (98),

$$\int |x'-x|^2 d\varrho \le CT \left(\int |\varpi|^2 + |\Xi| d\varrho \right) \le CT \left(T + \int \overline{L} d\varrho \right). \tag{103}$$

We now consider an additional small parameter $d > 0$. From (101), combined with (98),

$$\begin{aligned} \int |x'-x|^2 d\nu_T &\le C \left(\int \frac{|x'-x|^2}{d} + d|\varpi|^2 + d|\Xi|^p + d^{-1/(p-1)} d\varrho \right) \\ &\le C d^{-1/(p-1)} T + \frac{C}{d} \int |x'-x|^2 d\varrho + Cd \left(T + \int \overline{L} d\varrho \right). \end{aligned}$$

Hence, exploiting (103),

$$\int |x'-x|^2 d\nu_T \le C_d T + C(d+T) \int \overline{L} d\varrho \,, \tag{104}$$

where $C_d > 0$ may depend on d.
To get the desired estimates, it remains to estimate $\int \overline{L} d\varrho$. Fixed $x \in \mathbb{R}^n$, we take $\varpi_0 = \varpi_0(x)$, $\sigma_0 = \sigma_0(x)$ and $\Xi_0 = \Xi_0(x) = \sigma_0 \sigma_0^T$ in such a way that

$$\begin{aligned} &\sup_{\varpi, \Xi} -D_x \psi(x) \cdot \varpi - \Xi : D_x^2 \psi(x) - \overline{L}(x, \varpi, \Xi) \\ &\quad = -D_x \psi(x) \cdot \varpi_0 - \Xi_0 : D_x^2 \psi(x) - \overline{L}(x, \varpi_0, \Xi_0) \,. \end{aligned}$$

Indeed, note that, since $\overline{L}$ is coercive, by (98), and lower semicontinuous, by Proposition 14, we have that such ϖ_0 and Ξ_0 exist, although they may not be unique. Consider now the diffusion

$$d\mathbf{x}_0 = \mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(\mathbf{x}_0)dt + \mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(\mathbf{x}_0)dW_t\,, \tag{105}$$

with $\mathbf{x}_0(0) = x$ and let the measures $d\varrho_0(x',\omega,\Xi,s)$ be defined by

$$\begin{aligned}
&\int_{\mathbb{R}^n\times\mathbb{R}^n\times\mathrm{Mat}(n\times n)} \phi(x',\varpi,\Xi,s)d\varrho_0 \\
&:= E\int_0^T \phi(\mathbf{x}_0(s),\mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(\mathbf{x}_0(s)),\mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(\mathbf{x}_0(s)),s)ds
\end{aligned}$$

and $\nu_{0,T}(x')$ by

$$\int_{\mathbb{R}^n} \varphi(x')d\nu_{0,T}(x') := E\varphi(\mathbf{x}_0(T)).$$

Then, by Dynkin's Formula and the minimality of ϱ and ν_T,

$$\int \overline{L}d\varrho + \int \psi d\nu_T \le \int \overline{L}d\varrho_0 + \int \psi d\nu_{0,T}. \tag{106}$$

By Lemma 27, we have

$$\left|\int (x'-x)d\nu_{0,T}\right| \le CT \qquad \int |(x'-x)|^2\, d\nu_{0,T} \le CT.$$

Thus

$$\begin{aligned}
&\left|\int \psi(x')-\psi(x)\,d\nu_{0,T}(x')\right| \\
&\quad\le |D\psi(x)|\left|\int (x'-x)d\nu_{0,T}\right| + C\int |(x'-x)|^2\, d\nu_{0,T} \,\le\, CT\,.
\end{aligned} \tag{107}$$

On the other hand, by (102) and (104),

$$\begin{aligned}
&\left|\int \psi(x')-\psi(x)\,d\nu_T(x')\right| \\
&\quad\le |D\psi(x)|\left|\int (x'-x)d\nu_T\right| + C\int |(x'-x)|^2\, d\nu_T \\
&\quad\le CT + CT^{1/2}\left(\int \overline{L}d\varrho\right)^{1/2} + C_dT + C(d+T)\int \overline{L}d\varrho \\
&\quad\le C_dT + C(d+T)\int \overline{L}d\varrho\,.
\end{aligned} \tag{108}$$

Therefore, putting together (106), (107) and (108),

$$\begin{aligned}\int \overline{L} d\varrho &\le \left|\int \psi(x') - \psi(x) d\nu_{0,T}\right| + \left|\int \psi(x) - \psi(x') d\nu_T\right| \\ &\quad + \int \overline{L} d\varrho_0 \\ &\le C_d T + C(d+T) \int \overline{L} d\varrho.\end{aligned}$$

This yields the estimate in (97). Then, combining together (97), (102), (103) and (104), we obtain (94), (95) and (96). □

12. Viscosity Solutions

This last section is dedicated to the proof of the main result of the paper, that is that the homogenized limit is a viscosity solution to an Effective Hamilton-Jacobi equation.

Proposition 32. *There exists a function continuous function $\overline{H} : \mathrm{Mat}(n \times n) \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that $\bar{u}$ is a viscosity solution of*

$$-\bar{u}_t + \overline{H}\left(D_x^2 \bar{u}, D_x \bar{u}, x\right) = 0. \tag{109}$$

Proof. To prove that $\bar{u}$ is a viscosity solution, we are going to use the results from [Bit01]. To do so, we first define an operator:

$$\mathcal{T}_{t_1,t_2}\psi(x) := \inf_{\varrho,\nu_t} \int \overline{L} d\varrho + \int \psi d\nu_{t_2}, \tag{110}$$

where the infimum is taken over all measures ϱ on $\mathbb{R}^n \times \mathbb{R}^n \times [t_1, t_2]$ (or on $\mathbb{R}^n \times \mathbb{R}^n \times \mathrm{Mat}(n \times n) \times [t_1, t_2]$), and ν_t on $\mathbb{R}^n$ that satisfy (68) (or (69), according to case A, B or C) with T replaced by t_2 and t replaced by t_1, and $\nu_{t_1} = \delta(x)$.
We also set

$$\mathcal{T}_t := \mathcal{T}_{0,t}\,. \tag{111}$$

Let Y be the set of all $\psi \in C^2(\mathbb{R}^n)$ such that

$$\sup_{\epsilon>0} \sup_{x,\omega} \left| H\Big(D^2\psi, D\psi, x, \frac{x}{\epsilon}, \epsilon, \omega\Big)\right| < \infty. \tag{112}$$

Then, by (72) and Propositions 20, 24 and 26, we have that $\mathcal{T}_t$ maps Y into Lipschitz functions, for all $t \ge 0$. Define $X := \cup_{t\ge 0} \mathcal{T}_t Y$. If we prove that $\mathcal{T}_{t_1,t_2}$ satisfies the semigroup property: $\mathcal{T}_{t_1,t_2} \circ \mathcal{T}_{t_2,t_3} = \mathcal{T}_{t_1,t_3}$, for all $0 \le t_1 < t_2 < t_3$, it is then clear that X is an invariant set for $\mathcal{T}_t$. Note that Y contains $C_c^\infty(\mathbb{R}^n)$, thanks to (8), thus (H1) in [Bit01] holds.
To address the semigroup property, it suffices to observe that for $\epsilon > 0$ the evolution associated to the Hamilton-Jacobi equation $\mathcal{T}^\epsilon_{t_1,t_2}$ is a semigroup, which converges uniformly, and therefore the limit also satisfies the semigroup property. In further detail: one considers

$$\mathcal{T}^\epsilon_{t_1,t_2}\psi := \inf \int L d\mu^\epsilon + \int \psi d\nu^\epsilon_{t_2}\,,$$

where the infimum is taken on the measures ρ on $\mathbb{R}^n \times \mathbb{R}^n \times U \times [t_1, t_2]$ with total mass $t_2 - t_1$ and ν_t probability measures on $\mathbb{R}^n$ satisfying (30). From Proposition 9 and (5),

$$\mathcal{T}^{\epsilon}_{t_1,t_2}\psi = \inf E \int_{t_1}^{t_2} L\left(\mathbf{x}(s), \frac{\mathbf{x}(s)}{\epsilon}, \mathbf{v}(s), \omega\right) ds + \psi\left(\mathbf{x}(t_2)\right),$$

where the infimum above is taken over all trajectories of the controlled dynamics (3) with $\mathbf{x}(t_1) = x$. Then, $\mathcal{T}^{\epsilon}_{t_1,t_2}$ is a semigroup, by the optimality of the latter dynamic programming principle, thence so is $\mathcal{T}_{t_1,t_2}$, due to Propositions 24 and 26, as desired.
Let now the translation operator τ be defined by $\tau_y\psi(x) = \psi(x+y)$, for any $\psi : \mathbb{R}^n \to \mathbb{R}$. The hypothesis (H2) in [Bit01] requires that for all $f \in X$ and all $y \in \mathbb{R}^n$, $\tau_y f \in X$. However, this is not really necessary, and we may replace it by requiring for all $y \in \mathbb{R}^n$ and $\psi \in Y$, $\tau_y\psi \in Y$. This condition is then verified, thanks to (8).
Hypothesis (H3) of [Bit01] is obviously fulfilled here too.
Moreover, clearly, $\mathcal{T}_{t_1,t_2}$ is monotone, that is, if $\bar{\psi}_1 \leq \bar{\psi}_2$ then $\mathcal{T}_{t_1,t_2}\bar{\psi}_1 \leq \mathcal{T}_{t_1,t_2}\bar{\psi}_2$. Therefore, to establish the result it suffices to check that the continuity and regularity hypothesis I-IV that we spell next. We will need the following notation: for any sequence $d = (d_k)$ of positive reals we define

$$Q_d = \{\eta \in C^{\infty}_c(\mathbb{R}^n), \|D^{\alpha}\eta\|_{\infty} \leq d_k, |\alpha| \leq k\}.$$

I Continuity: for every $\psi \in X$ the function $(t, x) \mapsto \mathcal{T}_t[\psi](x)$ is continuous and for all $b > a \geq 0$ there exists $C = C(a, b, \psi)$ such that
$$|\mathcal{T}_t\psi| \leq C,$$
for any $t \in [a, b]$.

II Locality: for every $\psi_1, \psi_2 \in C^{\infty}(\mathbb{R}^n) \cap X$ and any fixed $x \in \mathbb{R}^n$, and $r > 0$, such that $\psi_1 = \psi_2$ in the ball $B(x, r)$ then
$$\mathcal{T}_{t-h,t}\psi_1 - \mathcal{T}_{t-h,t}\psi_2 = o(h),$$
as $h \to 0^+$.

III Regularity: for any sequence of positive numbers $d = (d_k)$, any compact set $K \subset \mathbb{R}^n$ and for every $\psi \in C^{\infty}(\mathbb{R}^n) \cap X$ there exists a function $m_{K,f,d}(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ such that $m_{K,f,d}(0^+) = 0$,
$$|\mathcal{T}_t[\psi + \lambda\eta] - \mathcal{T}_t[\psi] - \lambda\eta(x)| \leq m_{K,f,d}(\lambda)t,$$
for any $(x, \eta) \in K \times Q_d$ and any $\lambda, t \geq 0$.

IV Translation: for any compact subset $K \subset \mathbb{R}^n$ and every $\psi \in C^{\infty}_c(K)$, there exists a function $n_{K,\psi} : \mathbb{R}^+ \to \mathbb{R}^+$, with $n_{K,\psi}(0^+) = 0$ such that
$$|\tau_y\mathcal{T}_t[\psi](x) - \mathcal{T}_t[\tau_y\psi](x)| \leq n_{K,\psi}(|y|)t,$$
for any $x \in K$ and $t \geq 0$.

Property [I] holds by (74). Properties [II] and [III] follow from improved moment estimates of the previous section, as we explain next. Property [IV] follows from (72) and (83).
Let us now give the details about how to check Properties [II] and [III]. First, we will use Proposition 30 to prove that [II] holds. Assume that ψ_1 and ψ_2 satisfy the

conditions of [II], and let $\mathbf{x}$ be an optimal trajectory for ψ_2, that is $\mathbf{x}$ optimizes (5) for $\psi := \psi_2$ and $\mathbf{x}(t-h) = x$. Then, by (72) and Propositions 24 and 26,

$$\mathcal{T}_{t-h,t}\psi_2 = \lim_{\epsilon\to 0} E\left[\int_t^T L\left(\mathbf{x}(s), \frac{\mathbf{x}(s)}{\epsilon}, \mathbf{v}(s), \omega\right) ds + \psi_2(\mathbf{x}(T))\right]$$

while

$$\mathcal{T}_{t-h,t}\psi_1 \le \lim_{\epsilon\to 0} E\left[\int_t^T L\left(\mathbf{x}(s), \frac{\mathbf{x}(s)}{\epsilon}, \mathbf{v}(s), \omega\right) ds + \psi_2(\mathbf{x}(T))\right],$$

via a suitable subsequence.
Consequently, using that ψ_1 and ψ_2 agree on $B(x,r)$, we gather that

$$\mathcal{T}_{t-h,t}\psi_1 - \mathcal{T}_{t-h,t}\psi_2 \le (\|\psi_1\|_\infty + \|\psi_2\|_\infty)\, P(|\mathbf{x}-x| > r),$$

which is $o(h)$ by Proposition 30.
To prove [III], we argue in an analogous way. Namely, we take an optimal trajectory for ψ and we take limit for $\epsilon \to 0$, as above. In this case, we also exploit Lemma 28 to conclude that

$$\begin{aligned}
&\mathcal{T}_{T-h,T}(\psi+\lambda\eta) - \mathcal{T}_{T-h,T}\psi - \lambda\eta \\
&\le \lim_{\epsilon\to 0} E\psi(\mathbf{x}(T)) + \lambda\eta(\mathbf{x}(T)) - \psi(\mathbf{x}(T)) - \lambda\eta(\mathbf{x}(T-h)) \\
&\le \lim_{\epsilon\to 0} \lambda E(\eta(\mathbf{x}(T)) - \eta(\mathbf{x}(T-h))) \le C\|D\eta\|_\infty \lambda h \le C d_1 \lambda h,
\end{aligned}$$

as required. The desired result then follows from Theorem 3.1 of [Bit01]. □

We remark that Theorem 3.1 of [Bit01] explicitly gives that

$$-\overline{H}(D^2\psi(x), D\psi(x), x) = \lim_{T\to 0} \frac{\mathcal{T}_{0,T}\psi - \psi}{T}, \tag{113}$$

in the notation of (110) and (111).
Therefore, under the additional Assumption 16, we can characterize completely $\overline{H}$ as the Legendre transform of $\overline{L}$.

Proposition 33. *If also Assumption 16 holds, then $\overline{H}$ is given by*

$$\overline{H}(M,p,x) = \sup_{\Xi=\sigma\sigma^T, \sigma\in \mathrm{Mat}(n\times n), \varpi\in\mathbb{R}^n} -p\cdot\varpi - \Xi : M - \overline{L}(x,\varpi,\Xi). \tag{114}$$

Proof. We will be using formula (113) to establish the result. We only consider case C, since cases A and B are analogous, with just less variables involved.
By taking $\varphi(x,t) := t$ and $\varphi(x,t) := 1$ in (69), we see that if ϱ and $\bar\nu_T$ satisfy (69), then they have total mass T and 1, respectively. Also, the supremum in (114) is finite and attained, due to Propositions 14 and 15. Thus, take $\varpi_0 = \varpi_0(x)$, $\sigma_0 = \sigma_0(x)$ and $\Xi_0 = \Xi_0(x) = \sigma_0\sigma_0^T$ that satisfy

$$\begin{aligned}
&\sup_{\varpi,\Xi} -D_x\psi(x)\cdot\varpi - \Xi : D_x^2\psi(x) - \overline{L}(x,\varpi,\Xi) \\
&\quad = -D_x\psi(x)\cdot\varpi_0 - \Xi_0 : D_x^2\psi(x) - \overline{L}(x,\varpi_0,\Xi_0).
\end{aligned}$$

We also consider the dynamics of the associated control problem

$$d\mathbf{x}_0 = \mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(\mathbf{x}_0)dt + \mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(\mathbf{x}_0)dW_t\,,$$

and corresponding measures $\varrho_0(x',\omega,\Xi,s)$ and $\nu_{0,T}(x')$, given by

$$\begin{aligned}&\int_{\mathbb{R}^n\times\mathbb{R}^n\times\mathrm{Mat}(n\times n)} \phi(x',\varpi,\Xi,s)d\varrho_0\\ &:= E\int_0^T \phi(\mathbf{x}_0(s),\mathcal{S}^{\varpi}_{x_0,\varpi_0,\sigma_0}(\mathbf{x}_0(s)),\mathcal{S}^{\Xi}_{x_0,\varpi_0,\Xi_0}(\mathbf{x}_0(s)),s)ds\end{aligned}$$

and

$$\int_{\mathbb{R}^n}\varphi(x')d\nu_{0,T}(x') := E\varphi(\mathbf{x}_0(T)).$$

We have that (69) holds, by Dynkin's Formula, and so

$$\begin{aligned}\psi - \mathcal{T}_T\psi &\geq \psi - \int \psi d\bar{\nu}_{0,T} - \int \overline{L}d\varrho_0\\ &= \int -D_x\psi\cdot\varpi - \Xi : D_x^2\psi - \overline{L}d\varrho_0\\ &= \int -D_x\psi(x)\cdot\varpi_0 - \Xi_0 : D_x^2\psi(x) - \overline{L}(x,\varpi_0,\Xi_0)d\varrho_0\\ &\quad + \int \Big| D_x\psi\cdot\varpi + \Xi : D_x^2\psi\\ &\qquad + \overline{L} - D_x\psi(x)\cdot\varpi_0 - \Xi_0 : D_x^2\psi(x) - \overline{L}(x,\varpi_0,\Xi_0)\Big| d\varrho_0.\\ &\geq T\sup_{\varpi,\Xi} -D_x\psi(x)\cdot\varpi - \Xi : D_x^2\psi(x) - \overline{L}(x,\varpi,\Xi)\\ &\quad -C\int_0^T E|\mathbf{x}_0(s)-x|\,ds,\end{aligned}$$

thanks to Assumption 16.
That is,

$$\begin{aligned}\frac{\psi-\mathcal{T}_T\psi}{T} &\geq \sup_{\varpi,\Xi} -D_x\psi(x)\cdot\varpi - \Xi : D_x^2\psi(x) - \overline{L}(x,\varpi,\Xi)\\ &\quad - \frac{C}{T}\int_0^T E|\mathbf{x}_0(s)-x|\,ds\,.\end{aligned}\tag{115}$$

We now observe that, by Lemma 27,

$$\int_0^T E|\mathbf{x}_0(s)-x|\,ds \leq \int_0^T (E|\mathbf{x}_0(s)-x|^2)^{1/2}\,ds \leq CT^{3/2},$$

and so

$$\frac{1}{T}\int_0^T E|\mathbf{x}_0(s)-x|\,ds \to 0,\tag{116}$$

as $T\to 0$.

Consequently, making use of (113), (115) and (116), we thus deduce that

$$\overline{H}(D^2\psi(x), D\psi(x), x) \geq \sup_{\varpi,\Xi} -D_x\psi(x)\cdot\varpi - \Xi : D^2_x\psi(x) - \overline{L}(x,\varpi,\Xi). \tag{117}$$

Now we will establish the reverse inequality and complete the proof of (114). For that matter, let $\varrho(x',\varpi,\Xi,s)$ and $\nu_T(x')$ be minimizing measures for (73) under the constraint in (69), with $t=0$. Then,

$$\begin{aligned}
&T\sup_{\varpi,\Xi} -D_x\phi(x)\cdot\varpi - \Xi : D^2_x\phi(x) - \overline{L}(x,\varpi,\Xi)\\
\geq&\int -D_x\phi(x)\cdot\mathcal{S}^{\varpi}_{x',\varpi,\Xi}(x) - \mathcal{S}^{\Xi}_{x',\varpi,\Xi}(x) : D^2_x\phi(x) - \overline{L}(x,\mathcal{S}^{\varpi}_{x',\varpi,\Xi}(x),\mathcal{S}^{\Xi}_{x',\varpi,\Xi}(x))d\varrho\\
\geq&\int -D_x\phi(x)\cdot\varpi - \Xi : D^2_x\phi(x) - \overline{L}(x,\varpi,\Xi)\,d\varrho - C\int |x'-x|d\varrho\,,
\end{aligned}$$

thanks to Assumption 16.
The above estimate, the holonomy constraint and Proposition 31 then give

$$\begin{aligned}
&\sup_{\varpi,\Xi} -D_x\phi(x)\cdot\varpi - \Xi : D^2_x\phi(x) - \overline{L}(x,\varpi,\Xi)\\
\geq&\frac{\psi - \int\psi d\bar\nu_T - \int\overline{L}d\varrho}{T} - \frac{C}{T}\left(T\int|x'-x|^2d\varrho\right)^{1/2}\\
\geq&\frac{\psi - \mathcal{T}_T\psi}{T} - CT^{1/2}
\end{aligned}$$

as long as T is small enough.
Therefore, by sending T to zero and recalling (113),

$$\sup_{\varpi,\Xi} -D_x\psi(x)\cdot\varpi - \Xi : D^2_x\psi(x) - \overline{L}(x,\varpi,\Xi) \geq \overline{H}(D^2\psi(x), D\psi(x), x) - 0.$$

This and (117) complete the proof of (114). □

13. Conclusion

Proof of Theorem 1. Propositions 4 and 9 imply the first claim of Theorem 1.
The second claim follows from Propositions 18 and 20 and Remark 19.
By collecting the results in Propositions 24, 26 and 32, we obtain the last claim of Theorem 1. □

References

[AB02] Olivier Alvarez and Martino Bardi. Viscosity solutions methods for singular perturbations in deterministic and stochastic control. *SIAM J. Control Optim.*, 40(4):1159–1188 (electronic), 2001/02.

[Alv99] O. Alvarez. Homogenization of Hamilton-Jacobi equations in perforated sets. *J. Differential Equations*, 159(2):543–577, 1999.

[Att84] H. Attouch. *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[Bil95] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[Bit01] Samuel Biton. Nonlinear monotone semigroups and viscosity solutions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(3):383–402, 2001.

[BLP76] Alain Bensoussan, Jacques-Louis Lions, and Georges Papanicolaou. Homogénéisation, correcteurs et problèmes non-linéaires. *C. R. Acad. Sci. Paris Sér. A-B*, 282(22):Aii, A1277–A1282, 1976.

[CC95] Luis A. Caffarelli and Xavier Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[CDI01] I. Capuzzo-Dolcetta and H. Ishii. On the rate of convergence in homogenization of Hamilton-Jacobi equations. *Indiana Univ. Math. J.*, 50(3):1113–1129, 2001.

[Con96] Marie C. Concordel. Periodic homogenization of Hamilton-Jacobi equations: additive eigenvalues and variational formula. *Indiana Univ. Math. J.*, 45(4):1095–1117, 1996.

[CS02] Moon Jung Cho and Richard H. Stockbridge. Linear programming formulation for optimal stopping problems. *SIAM J. Control Optim.*, 40(6):1965–1982 (electronic), 2002.

[DiB95] Emmanuele DiBenedetto. *Partial differential equations*. Birkhäuser Boston Inc., Boston, MA, 1995.

[EG01] L. C. Evans and D. Gomes. Effective Hamiltonians and averaging for Hamiltonian dynamics. I. *Arch. Ration. Mech. Anal.*, 157(1):1–33, 2001.

[Eva82] Lawrence C. Evans. Classical solutions of fully nonlinear, convex, second-order elliptic equations. *Comm. Pure Appl. Math.*, 35(3):333–363, 1982.

[Eva90] Lawrence C. Evans. *Weak convergence methods for nonlinear partial differential equations*, volume 74 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1990.

[Eva92] Lawrence C. Evans. Periodic homogenisation of certain fully nonlinear partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(3-4):245–265, 1992.

[Eva97] Lawrence C. Evans. Regularity for fully nonlinear elliptic equations and motion by mean curvature. In *Viscosity solutions and applications (Montecatini Terme, 1995)*, volume 1660 of *Lecture Notes in Math.*, pages 98–133. Springer, Berlin, 1997.

[Eva98] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[Eva06a] Lawrence C. Evans. *An Introduction to Mathematical Optimal Control Theory*. 2006. Preprint Version 0.1. Available on-line at `http://math.berkeley.edu/~evans/`.

[Eva06b] Lawrence C. Evans. *An Introduction to Stochastic Differential Equations*. 2006. Preprint Version 1.2. Available on-line at `http://math.berkeley.edu/~evans/`.

[FR75] Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control*. Springer-Verlag, Berlin, 1975. Applications of Mathematics, No. 1.

[FS06] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[Gom03] Diogo Aguiar Gomes. Perturbation theory for viscosity solutions of Hamilton-Jacobi equations and stability of Aubry-Mather sets. *SIAM J. Math. Anal.*, 35(1):135–147 (electronic), 2003.

[Gom06] Diogo Aguiar Gomes. *Lecture Notes on Calculus of Variations*. 2006. Preliminary Version. Available on-line at `http://www.math.ist.utl.pt/~dgomes/`.

[GR06] Vladimir Gaitsgory and Sergey Rossomakhine. Linear programming approach to deterministic long run average problems of optimal control. *SIAM J. Control Optim.*, 44(6):2006–2037 (electronic), 2006.

[Ish99] Hitoshi Ishii. Homogenization of the Cauchy problem for Hamilton-Jacobi equations. In *Stochastic analysis, control, optimization and applications*, Systems Control Found. Appl., pages 305–324. Birkhäuser Boston, Boston, MA, 1999.

[Ish00] Hitoshi Ishii. Almost periodic homogenization of Hamilton-Jacobi equations. In *International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999)*, pages 600–605. World Sci. Publishing, River Edge, NJ, 2000.

[Kos06] Elena Kosygina. Homogenization of stochastic Hamilton-Jacobi equations: brief review of methods and applications. 2006. Preprint. Available on-line at `http://faculty.baruch.cuny.edu/ekosygina/review.pdf`.

[KRV06] Elena Kosygina, Fraydoun Rezakhanlou, and Srinivasa R. S. Varadhan. Stochastic homogenization of Hamilton-Jacobi-Bellman equations. *Comm. Pure Appl. Math.*, 2006. To appear. Available on-line at `http://faculty.baruch.cuny.edu/ekosygina/CPAMHJ.pdf`.

[Kry80] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York, 1980. Translated from the Russian by A. B. Aries.

[KS] T. Kurtz and Richard H. Stockbridge. Linear programming formulation of stochastic control problems. *Proc. 37th IEEE - CDC.*

[LPV86] Pierre-Louis Lions, George Papanicolaou, and Srinivasa R. S. Varadhan. Homogenization of Hamilton-Jacobi equations. 1986. Unpublished paper.

[LS03] Pierre-Louis Lions and Panagiotis E. Souganidis. Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting. *Comm. Pure Appl. Math.*, 56(10):1501–1524, 2003.

[LS05] Pierre-Louis Lions and Panagiotis E. Souganidis. Homogenization of degenerate second-order PDE in periodic and almost periodic environments and applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(5):667–677, 2005.

[MS98] Marta Susana Mendiondo and Richard H. Stockbridge. Approximation of infinite-dimensional linear programming problems which arise in stochastic control. *SIAM J. Control Optim.*, 36(4):1448–1472 (electronic), 1998.

[RT00] Fraydoun Rezakhanlou and James E. Tarver. Homogenization for stochastic Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.*, 151(4):277–309, 2000.

[Shu92] M. A. Shubin, editor. *Partial differential equations. I*, volume 30 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1992. Foundations of the classical theory, A translation of Differentsialnye uravneniya s chastnymi proizvodnymi, 1, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1988, Translation by R. Cooke, Translation edited by Yu. V. Egorov and M. A. Shubin.

[Sou99] Panagiotis E. Souganidis. Stochastic homogenization of Hamilton-Jacobi equations and some applications. *Asymptot. Anal.*, 20(1):1–11, 1999.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Wan90] Lihe Wang. On the regularity theory of fully nonlinear parabolic equations. *Bull. Amer. Math. Soc. (N.S.)*, 22(1):107–114, 1990.

[Wan92a] Lihe Wang. On the regularity theory of fully nonlinear parabolic equations. I. *Comm. Pure Appl. Math.*, 45(1):27–76, 1992.

[Wan92b] Lihe Wang. On the regularity theory of fully nonlinear parabolic equations. II. *Comm. Pure Appl. Math.*, 45(2):141–178, 1992.

[Wan92c] Lihe Wang. On the regularity theory of fully nonlinear parabolic equations. III. *Comm. Pure Appl. Math.*, 45(3):255–262, 1992.

Diogo A. Gomes
Departamento de Matemática, Instituto Superior Técnico, Lisboa, 1049-001, Portugal
e-mail: dgomes@math.ist.utl.pt

Enrico Valdinoci
Dipartimento di Matematica, Università di Roma Tor Vergata, Roma, I-00133, Italy
e-mail: valdinoci@mat.uniroma2.it