Stochastic Homogenization of Hamilton-Jacobi-Bellman Equations

ELENA KOSYGINA
Baruch College

FRAYDOUN REZAKHANLOU
University of California, Berkeley

AND

S. R. S. VARADHAN
Courant Institute

Abstract

We study the homogenization of some Hamilton-Jacobi-Bellman equations with a vanishing second-order term in a stationary ergodic random medium under the hyperbolic scaling of time and space. Imposing certain convexity, growth, and regularity assumptions on the Hamiltonian, we show the locally uniform convergence of solutions of such equations to the solution of a deterministic "effective" first-order Hamilton-Jacobi equation. The effective Hamiltonian is obtained from the original stochastic Hamiltonian by a minimax formula. Our homogenization results have a large-deviations interpretation for a diffusion in a random environment. © 2005 Wiley Periodicals, Inc.

1 Introduction

Homogenization of a differential equation is the process of replacing rapidly varying coefficients (or functions) by new ones such that the solutions are close. A simple illustration is provided by the following example: Suppose that $a(x)$ is a periodic or stationary random function on $\mathbb{R}$ satisfying $a \leqslant a(x) \leqslant b$ for some constants $0 < a < b < \infty$. For $\varepsilon > 0$ one can consider the elliptic operator

$$\mathcal{L}_\varepsilon = \frac{d}{dx} a\left(\frac{x}{\varepsilon}\right) \frac{d}{dx}.$$

For small ε, $\mathcal{L}_\varepsilon$ can be replaced by

$$\mathcal{L} = \frac{d}{dx} \bar{a} \frac{d}{dx},$$

where $\bar{a} = [\mathbb{E}[1/a]]^{-1}$ is the "harmonic mean" of $a(\cdot)$. $\mathbb{E}$ is the expectation in the random case or the average over a period in the periodic case. As $\varepsilon \to 0$, solutions

Communications on Pure and Applied Mathematics, Vol. LIX, 0001–0033 (2006)

of parabolic or elliptic equations of the form

$$\frac{du}{dt} = \mathcal{L}_\varepsilon u, \quad u(0,x) = f(x),$$

or

$$\lambda u - \mathcal{L}_\varepsilon u = f$$

are close to the corresponding solutions with $\mathcal{L}_\varepsilon$ replaced by $\mathcal{L}$.

There is a vast literature on periodic and quasi-periodic homogenization of linear and nonlinear partial differential equations (see, for example, monographs [1, 2, 5, 7, 10, 15] and references therein). Many results obtained in the periodic setting are based on the existence of "correctors," i.e., periodic solutions to an auxiliary problem, which arise from a formal expansion in a scaling parameter ε. In connection with this approach for Hamilton-Jacobi type and more general equations, let us mention just several representative works [4, 6, 8, 11], where one can find further references. It turns out [12] that for Hamilton-Jacobi equations with stationary ergodic Hamiltonians, correctors may not exist in general. Homogenization results for stationary ergodic environments are usually obtained by an application of a version of the ergodic theorem (see, for instance, [3, 10, 16, 17, 19]).

In the current work we will be interested in solutions to certain Hamilton-Jacobi-Bellman equations. These are solutions u_ε to equations of the type

$$\frac{\partial u_\varepsilon}{\partial t} = \frac{\varepsilon}{2}\Delta u_\varepsilon + H\left(\nabla u_\varepsilon, \frac{x}{\varepsilon}, \omega\right), \quad (t,x) \in [0,\infty) \times \mathbb{R}^d, \tag{1.1}$$

with initial condition $u_\varepsilon(0,x) = f(x)$. Here $H(p,x,\omega)$ is a convex function of p that is a stationary random process in x. We are interested in establishing a result of the form $u_\varepsilon \to u$ as $\varepsilon \to 0$ where u is the solution of

$$\frac{\partial u}{\partial t} = \overline{H}(\nabla u), \quad u(0,x) = f(x). \tag{1.2}$$

The solutions u_ε and u have natural variational representations and we will consider these problems from this perspective. Therefore, it will not be necessary for us to characterize these solutions as viscosity solutions.

These types of problems arise naturally if we consider questions of large deviations for diffusions in a random environment. More specifically, consider a diffusion process on $\mathbb{R}^d$ with generator

$$\mathcal{A}_\omega = \frac{1}{2}\Delta + \langle b(x,\omega), \nabla\rangle,$$

i.e., Brownian motion with a random drift $b(x,\omega)$ that is assumed to be an ergodic stationary process in $x \in \mathbb{R}^d$. We are interested in considering the probabilities of large deviations of $x(t)/t$ as $t \to \infty$. This amounts to establishing the limiting behavior

$$\lim_{t\to\infty} \frac{1}{t} \log E^{Q_0^{b,\omega}}[\exp[\langle\theta, x(t)\rangle]] = \overline{H}(\theta).$$

Here $Q_x^{b,\omega}$ is the diffusion process corresponding to $\mathcal{A}_\omega$ starting from $x \in \mathbb{R}^d$ at time 0. The above limit is supposed to exist almost surely in ω. The probabilities of large deviations are now determined in the standard manner, the rate function being the conjugate function $\mathcal{I}(y)$ defined for $y \in \mathbb{R}^d$, by

$$\mathcal{I}(y) = \sup_{\theta \in \mathbb{R}^d} [\langle \theta, y \rangle - \overline{H}(\theta)]. \tag{1.3}$$

The quantity

$$v(t, x) = E^{Q_x^{b,\omega}}[\exp\langle \theta, x(t) \rangle]$$

solves the equation

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v + \langle b(x, \omega), \nabla v \rangle, \quad v(0, x) = \exp\langle \theta, x \rangle.$$

We are interested in the behavior of $\frac{1}{t}\log v(t, 0)$ as $t \to \infty$. If we define

$$u_\varepsilon(t, x) = \varepsilon \log v\left(\frac{t}{\varepsilon}, \frac{x}{\varepsilon}\right),$$

then the problem reduces to the study of the behavior of $u_\varepsilon(1, 0)$ as $\varepsilon \to 0$. The equation satisfied by u_ε is

$$\frac{\partial u_\varepsilon}{\partial t} = \frac{\varepsilon}{2}\Delta u_\varepsilon + \frac{1}{2}\|\nabla u_\varepsilon\|^2 + \left\langle b\left(\frac{x}{\varepsilon}, \omega\right), \nabla u_\varepsilon \right\rangle, \quad u(0, x) = \langle \theta, x \rangle.$$

This is, of course, a special case of (1.1) with

$$H(p, x, \omega) = \frac{1}{2}\|p\|^2 + \langle b(x, \omega), p \rangle.$$

Remark 1.1. After our work was completed, we received a preprint [13] that deals with more general "viscous" Hamilton-Jacobi equations and in the special case that we consider overlaps considerably with our results. The main tools in [13] are the uniform gradient estimate on solutions u_ε and the subadditive ergodic theorem. The approach presented in our paper is different. It is based on a direct application of the standard ergodic theorem to obtain a uniform lower bound on solutions and on a straightforward construction of "supercorrectors" for an upper bound. The minimax formula for the effective Hamiltonian arises naturally in the proof. Another characteristic feature of our method is that it does not require a uniform gradient estimate of u_ε even for subquadratic Hamiltonians. The tradeoff is that we need a stronger regularity assumption on the Hamiltonian (see (H3)) in the next section. For instance, in the special case of

$$H(p, \omega) = a(\omega)\|p\|^\alpha$$

with $\alpha > 1$, in addition to uniform upper and lower bounds on a, we would need uniform continuity of $a(\tau_y\omega)$, whereas [13] would assume a uniform bound on the gradient in y of $a(\tau_y\omega)$. Moreover, if there is a way of getting uniform gradient bounds for perturbations of the form $H(p, \omega) + \delta\|p\|^k$ that are uniform in δ, we can easily choose $k > d$ and put ourselves in a situation where (H4) holds. It is not hard to let $\delta \to 0$ and recover Theorem 2.2.

2 Main Results

We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which $\mathbb{R}^d$ acts as a group $\{\tau_x : x \in \mathbb{R}^d\}$ of measure-preserving transformations. $\mathbb{P}$ is assumed to be ergodic under this action. Let the function $H(p, \omega) : \mathbb{R}^d \times \Omega \to \mathbb{R}$ be convex in p and have the following properties:

(H1) For all $p \in \mathbb{R}^d$, $\omega \in \Omega$,

$$c_1\|p\|^\alpha - c_2 \leqslant H(p, \omega) \leqslant c_3\|p\|^\beta + c_4 \tag{2.1}$$

for some positive constants c_1, c_2, c_3, and c_4, $1 < \alpha < \beta < \infty$. Equivalently, its convex conjugate $L(q, \omega)$, defined by

$$L(q, \omega) = \sup_p [\langle p, q\rangle - H(p, \omega)], \tag{2.2}$$

satisfies

$$c_1\|q\|^{\beta'} - c_2 \leqslant L(q, \omega) \leqslant c_3\|q\|^{\alpha'} + c_4 \tag{2.3}$$

with $\alpha' = \alpha/(\alpha - 1)$, $\beta' = \beta/(\beta - 1)$, and possibly a different set of constants.

(H2) $H(p, \tau_x\omega)$ is uniformly continuous in x; i.e., for any $\ell < \infty$,

$$\lim_{\delta\to 0} \sup_{\|x\|\leqslant\delta} \sup_{\|p\|\leqslant\ell} \sup_{\omega} |H(p, \tau_x\omega) - H(p, \omega)| = 0. \tag{2.4}$$

It is not difficult to see, by using relation (2.2) and the bounds (2.1), that $L(q, \omega)$ satisfies a similar estimate

$$\lim_{\delta\to 0} \sup_{\|x\|\leqslant\delta} \sup_{\|q\|\leqslant\ell} \sup_{\omega} |L(q, \tau_x\omega) - L(q, \omega)| = 0. \tag{2.5}$$

For some of the results we will need stronger regularity or convexity assumptions.

(H3) There exists a function $\nu(\delta) \to 0$ as $\delta \to 0$ and $C > 0$ such that for $\|x\| \leqslant \delta$ and $\omega \in \Omega$

$$L(q, \tau_x\omega) \leqslant (1 + \nu(\delta))L(q, \omega) + C\nu(\delta). \tag{2.6}$$

Equivalently, there exists a function $\nu(\delta) \to 0$ as $\delta \to 0$ and $C > 0$ such that for $\|x\| \leqslant \delta$ and $\omega \in \Omega$

$$H(p, \tau_x\omega) \geqslant (1 + \nu(\delta))H((1 + \nu(\delta))^{-1}p, \omega) - C\nu(\delta). \tag{2.7}$$

(H4) Assumption (H1) holds with $\alpha > d$.

(H5) Assumption (H1) holds with $\alpha > 2$ and there exist functions $\gamma(\delta) > 0$ and $C(\delta) \to 0$ as $\delta \to 0$ such that for all $\omega \in \Omega$

$$(1-\delta)H((1-\delta)^{-1}p, \omega) \geqslant H(p, \omega) + \gamma(\delta)\|p\|^2 - C(\delta). \tag{2.8}$$

For (2.8) to hold, it is sufficient that (uniformly in ω) $D^2H(p, \omega) \geqslant cI$ on the set $\{p \in \mathbb{R}^d : \|p\| \geqslant k\}$ for some $c, k > 0$. Here I is the identity matrix.

For any given $\varepsilon > 0$ and $\omega \in \Omega$, we consider the solution $u_\varepsilon = u_\varepsilon(t, x, \omega)$ of equation (1.1), which we write as

$$\frac{\partial u_\varepsilon}{\partial t} = \frac{\varepsilon}{2}\Delta u_\varepsilon + H(\nabla u_\varepsilon, \tau_{x/\varepsilon}\omega), \quad (t, x) \in [0, \infty) \times \mathbb{R}^d, \tag{2.9}$$

with the initial condition $u_\varepsilon(0, x) = f(x)$. We assume that f satisfies the following condition:

(H6) The function f is uniformly continuous on $\mathbb{R}^d$. This implies that for every $\delta > 0$ there is a constant $K_\delta > 0$ such that for all $x, y \in \mathbb{R}^d$

$$|f(x) - f(y)| \leqslant K_\delta\|x - y\| + \delta. \tag{2.10}$$

We note that the solution $u_\varepsilon(t, x, \omega)$ of (2.9) is equal to $\varepsilon v_\varepsilon(t/\varepsilon, x/\varepsilon, \omega)$, the rescaled version of v_ε that solves

$$\frac{\partial v_\varepsilon}{\partial t} = \frac{1}{2}\Delta v_\varepsilon + H(\nabla v_\varepsilon, \tau_x\omega), \quad (t, x) \in [0, \infty) \times \mathbb{R}^d, \tag{2.11}$$

with $v_\varepsilon(0, x) = \varepsilon^{-1}f(\varepsilon x)$.

We now construct the convex function $\overline{H}(p)$ that appears in (1.2). The translation group $\{\tau_x : x \in \mathbb{R}^d\}$ acting on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ will have infinitesimal generators $\{\nabla_i : 1 \leqslant i \leqslant d\}$ in the coordinate directions and the corresponding Laplace operator $\Delta = \sum_i \nabla_i^2$. For reasonable choices of $b(\omega) : \Omega \to \mathbb{R}^d$, the operator

$$\mathcal{A}_b = \frac{1}{2}\Delta + \langle b(\omega), \nabla\rangle$$

will define a Markov process on Ω. Construction of this Markov process is not difficult. Given a starting point $\omega \in \Omega$, we define $b(x, \omega) : \mathbb{R}^d \to \mathbb{R}^d$ by $b(x, \omega) = b(\tau_x\omega)$. This allows us to define the diffusion $Q_0^{b,\omega}$, starting from 0 at time 0, in the random environment that corresponds to the generator

$$\frac{1}{2}\Delta + \langle b(x, \omega), \nabla\rangle.$$

The diffusion is then lifted to Ω by evolving ω randomly in time by the rule $\omega(t) = \tau_{x(t)}\omega$. The induced measure $P^{b,\omega}$ defines the Markov process on Ω that corresponds to $\mathcal{A}_b$. The problem of finding the invariant measures for the process $P^{b,\omega}$ with generator $\mathcal{A}_b$ on Ω is very hard and nearly impossible to solve. However, if we can find a density $\varphi > 0$ such that $\varphi\, d\mathbb{P}$ is an invariant ergodic probability measure for $\mathcal{A}_b$, then we have by the ergodic theorem

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t F(\omega(s))ds = \int_\Omega F(\omega)\varphi(\omega)d\mathbb{P}$$

a.e. $P^{b,\omega}$ or in $L^1(P^{b,\omega})$ for almost all ω with respect to $\mathbb{P}$.

Let us denote by $\mathbf{B}$ the space of essentially bounded maps from $\Omega \to \mathbb{R}^d$ and by $\mathbf{D}$ the space of probability densities $\varphi : \Omega \to \mathbb{R}$ relative to $\mathbb{P}$, with φ, $\nabla\varphi$, and $\nabla^2\varphi$ essentially bounded and φ in addition having a positive essential lower bound. Let us denote by $\mathcal{E}$ the following subset of $\mathbf{B} \times \mathbf{D}$:

$$\mathcal{E} = \Big\{(b, \varphi) \in \mathbf{B} \times \mathbf{D} : \tfrac{1}{2}\Delta\varphi = \nabla \cdot (b\varphi)\Big\}. \tag{2.12}$$

Here we assume that the equation $\frac{1}{2}\Delta\varphi = \nabla \cdot (b\varphi)$ is satisfied in the weak sense. We define the convex function $\overline{H}$ on $\mathbb{R}^d$ by

$$\overline{H}(p) = \sup_{(b,\varphi)\in\mathcal{E}} \Big[\langle p, \mathbb{E}[b(\omega)\varphi(\omega)]\rangle - \mathbb{E}[L(b(\omega), \omega)\varphi(\omega)]\Big], \tag{2.13}$$

where $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$. The corresponding variational solution of (1.2) is given by

$$u(t, x) = \sup_y \left(f(y) - t\mathcal{I}\left(\frac{y - x}{t}\right)\right), \tag{2.14}$$

where $\mathcal{I}$ is related to $\overline{H}$ by the duality relation (1.3). Our main result is stated below.

THEOREM 2.1 (Lower Bound) *Assume* (H1)*,* (H2)*, and* (H6)*. Let $u_\varepsilon(t, x, \omega)$ be the solution of* (2.9) *and $u(t, x)$ be given by* (2.14)*. Then with probability* 1 *for any $\ell, T > 0$, we have*

$$\liminf_{\varepsilon\to 0} \inf_{0\leqslant t\leqslant T} \inf_{\|x\|\leqslant \ell} (u_\varepsilon(t, x, \omega) - u(t, x)) \geqslant 0.$$

THEOREM 2.2 (Upper Bound) *Assume* (H1)*,* (H2)*,* (H6)*, and either* (H3)*,* (H4)*, or* (H5)*. Let $u_\varepsilon(t, x, \omega)$ be the solution of* (2.9) *and $u(t, x)$ be given by* (2.14)*. Then with probability* 1 *for any $\ell, T > 0$, we have*

$$\limsup_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} \sup_{\|x\|\leqslant \ell} (u_\varepsilon(t, x, \omega) - u(t, x)) \leqslant 0.$$

Theorems 2.1 and 2.2 immediately imply the homogenization result under appropriate conditions.

THEOREM 2.3 *Assume* (H1)*,* (H2)*,* (H6)*, and either* (H3)*,* (H4)*, or* (H5)*. Then with probability* 1

$$\lim_{\varepsilon\to 0} |u_\varepsilon(t, x, \omega) - u(t, x)| = 0 \tag{2.15}$$

locally uniformly in $(t, x) \in [0, \infty) \times \mathbb{R}^d$.

Remark 2.4. It is easy to see from the definition of $\overline{H}(p)$ that it satisfies the same upper and lower bounds given in (2.1). We can always choose $b(\omega) = b$ a constant vector $b \in \mathbb{R}^d$, and $\mathbb{P}$ is still an ergodic invariant measure for the Markov process with generator $\mathcal{A}_b$ on Ω. Therefore, using the dual upper bound on L, which is uniform in ω, we get

$$\overline{H}(p) \geqslant \sup_{b \in \mathbb{R}^d} [\langle p, b\rangle - \mathbb{E}[L(b,\omega)]] \geqslant c_1 \|p\|^\alpha - c_2. \tag{2.16}$$

It is just as easy to note that

$$\overline{H}(p) \leqslant \sup_{b \in \mathbb{R}^d} [\langle p, b\rangle - \inf_\omega L(b,\omega)] \leqslant c_3 \|p\|^\beta + c_4. \tag{2.17}$$

3 Outline of Proof

The first step in establishing the lower bound is the variational representation of solutions of Hamilton-Jacobi-Bellman equations (see [9]). Let $\mathcal{C}$ be the set of all bounded maps $c(s,x)$ from $[0,T] \times \mathbb{R}^d$ to $\mathbb{R}^d$ such that $\sup_{s,x} \|c(s,x)\| < \infty$. Consider the diffusion Q_x^c on $\mathbb{R}^d$ starting from $x \in \mathbb{R}^d$ at time 0 with time-dependent generator

$$\frac{1}{2}\Delta + c(s,x) \cdot \nabla$$

in the time interval $[0,t]$. For each $c \in \mathcal{C}$ and $\omega \in \Omega$ we consider

$$v_c(t,x,\omega) = E^{Q_x^c}\left(f(x(t)) - \int_0^t L\big(c(s,x(s)), \tau_{x(s)}\omega\big) ds \right),$$

where $L(c,\omega)$ is as in (2.2). If

$$v(t,x,\omega) = \sup_{c \in \mathcal{C}} v_c(t,x,\omega),$$

then v is the solution of

$$\frac{\partial v}{\partial t} = \frac{1}{2}\Delta v + H(\nabla v, \tau_x \omega) \tag{3.1}$$

with $v(0,x) = f(x)$.

There is a simple relation between $v(t,y,\cdot)$ and $v(t,0,\cdot)$. If we define $f^y(x) = f(x+y)$, then the solution of (3.1) with initial data $v(0,x) = f^y(x)$ and $\omega' = \tau_y \omega$ is given by

$$v^y(t,x,\omega') = v^y(t,x,\tau_y\omega) = v(t,x+y,\omega). \tag{3.2}$$

In particular,

$$v(t,y,\omega) = v^y(t,0,\tau_y\omega).$$

The solution u_ε of (2.9) with initial data $f(x)$ is related to the solution v_ε of (3.1) with initial data $\varepsilon^{-1} f(\varepsilon x)$ by

$$u_\varepsilon(t,x,\omega) = \varepsilon v_\varepsilon(\varepsilon^{-1}t, \varepsilon^{-1}x, \omega). \tag{3.3}$$

Relation (3.2) translates to

$$u_\varepsilon^y(t,x,\omega) = u_\varepsilon(t, x+y, \tau_{-y/\varepsilon}\omega).$$

We therefore obtain the following variational expression for $u_\varepsilon(t,x)$:

$$\begin{aligned} u_\varepsilon(t,x,\omega) &= \sup_{c\in\mathcal{C}} E^{Q^c_{x/\varepsilon}}\left(f(\varepsilon x(\varepsilon^{-1}t)) - \varepsilon \int_0^{t/\varepsilon} L\big(c(s,x(s)), \tau_{x(s)}\omega\big) ds \right) \\ &= \sup_{c\in\mathcal{C}} E^{Q^{\varepsilon,c}_x}[f(x(t)) - \xi_\varepsilon(t)] \end{aligned} \tag{3.4}$$

where $Q^{\varepsilon,c}_x$ is the diffusion on $\mathbb{R}^d$ starting from x corresponding to the generator

$$\frac{\varepsilon}{2}\Delta + c(\varepsilon^{-1}s, \varepsilon^{-1}x)\cdot\nabla;$$

i.e., almost surely with respect to $Q^{\varepsilon,c}_x$, $x(t)$ satisfies

$$x(t) = x + \int_0^t c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))ds + \sqrt{\varepsilon}\beta(t) \tag{3.5}$$

and

$$\xi_\varepsilon(t) = \int_0^t L\big(c(\varepsilon^{-1}s, \varepsilon^{-1}x(s)), \tau_{\varepsilon^{-1}x(s)}\omega\big)ds. \tag{3.6}$$

Since the supremum over $c\in\mathcal{C}$ is taken for each ω, one can choose c to depend on ω. A special choice for $c(t,x)$, one that depends on $\omega\in\Omega$ but not on t, is the choice $c(t,x) = c(t,x,\omega) = c(x,\omega) = b(\tau_x\omega)$ with $(b,\varphi)\in\mathcal{E}$. With that choice we can consider either the process $\{Q^{b,\omega}_x\}$ on $\mathbb{R}^d$ or the process $\{P^{b,\omega}\}$ with values in Ω. It is easy to see that for any $y\in\mathbb{R}^d$, the translation map $\hat{\tau}_y$ on $C([0,T];\mathbb{R}^d)$ defined by $x(\cdot)\to x(\cdot)+y$ has the property

$$Q^{b,\omega}_y = Q^{b,\tau_y\omega}_0 \hat{\tau}_y^{-1},$$

which is essentially a restatement of (3.2). Since $(b,\varphi)\in\mathcal{E}$, by the ergodic theorem we have

$$\lim_{\varepsilon\to 0} \varepsilon \int_0^{t/\varepsilon} b(\omega(s))ds = t\int b(\omega)\varphi(\omega)d\mathbb{P} \stackrel{\text{def}}{=} m(b,\varphi)t \tag{3.7}$$

and

$$\lim_{\varepsilon\to 0} \varepsilon \int_0^{t/\varepsilon} L\big(b(\omega(s)), \omega(s)\big)ds = t\int L(b(\omega),\omega)\varphi(\omega)d\mathbb{P} \stackrel{\text{def}}{=} h(b,\varphi)\,t. \tag{3.8}$$

Both limits are valid in $L^1(P^{b,\omega})$ for $\mathbb{P}$ almost all ω. If we define $\mathbf{A}\subset\mathbb{R}^d\times\mathbb{R}$ as

$$\mathbf{A} = \{(m(b,\varphi), h(b,\varphi)) : (b,\varphi)\in\mathcal{E}\}, \tag{3.9}$$

then

$$\liminf_{\varepsilon\to 0} u_\varepsilon(t,0,\omega) \geqslant [f(mt) - ht]$$

for every $(m, h) \in \mathbf{A}$. Therefore, for almost all ω with respect to $\mathbb{P}$

$$
\begin{aligned}
\liminf_{\varepsilon \to 0} u_\varepsilon(t, 0, \omega) &\geqslant \sup_{(m,h)\in\mathbf{A}} [f(mt) - ht] \\
&= \sup_{y\in\mathbb{R}^d} (f(y) - t\mathcal{I}(y/t)) = u(t, 0).
\end{aligned}
\tag{3.10}
$$

This is a very weak form of convergence, and work has to be done in order to strengthen it to locally uniform convergence.

The upper bound is first obtained for linear f and then extended to general f. By using the convex duality and the minimax theorem, the right-hand side of (3.10) is rewritten in terms of the dual problem. If we take $f(x) = \langle p, x\rangle$, we have established an asymptotic lower bound for u_ε, which is the solution

$$u(t, x) = \langle p, x\rangle + t\overline{H}(p)$$

of (1.2) with $u(0, x) = \langle p, x\rangle$. Here

$$
\begin{aligned}
\overline{H}(p) &= \sup_{(b,\varphi)\in\mathcal{E}} \mathbb{E}\big[[\langle p, b(\omega)\rangle - L(b(\omega), \omega)]\varphi(\omega)\big] \\
&= \sup_\varphi \sup_b \inf_\psi \mathbb{E}\big[[\langle p, b(\omega)\rangle - L(b(\omega), \omega) + \mathcal{A}_b\psi(\omega)]\varphi(\omega)\big] \\
&= \sup_\varphi \inf_\psi \sup_b \mathbb{E}\big[[\langle p, b(\omega)\rangle - L(b(\omega), \omega) + \mathcal{A}_b\psi(\omega)]\varphi(\omega)\big] \\
&= \sup_\varphi \inf_\psi \sup_b \mathbb{E}\big[[\langle p + \nabla\psi(\omega), b(\omega)\rangle - L(b(\omega), \omega) + \tfrac{1}{2}\Delta\psi(\omega)]\varphi(\omega)\big] \\
&= \sup_\varphi \inf_\psi \mathbb{E}\big[[H(p + \nabla\psi(\omega), \omega) + \tfrac{1}{2}\Delta\psi(\omega)]\varphi(\omega)\big] \\
&= \inf_\psi \sup_\varphi \mathbb{E}\big[[H(p + \nabla\psi(\omega), \omega) + \tfrac{1}{2}\Delta\psi(\omega)]\varphi(\omega)\big] \\
&= \inf_\psi \operatorname*{ess\,sup}_\omega \big[H(p + \nabla\psi(\omega), \omega) + \tfrac{1}{2}\Delta\psi(\omega)\big].
\end{aligned}
$$

We have used the fact that

$$\inf_\psi \mathbb{E}[\varphi\mathcal{A}_b\psi(\omega)] = -\infty$$

unless $\varphi\, d\mathbb{P}$ is an invariant measure for $\mathcal{A}_b$, in which case it is 0. It follows that for any $\delta > 0$, there exists a "ψ" such that

$$\frac{1}{2}\Delta\psi(\omega) + H(\theta + \nabla\psi(\omega), \omega) \leqslant \overline{H}(\theta) + \delta.$$

The "ψ" is a weak object, and one has to do some work before we can use it as a test function and obtain the upper bound by comparison. The interchange of inf and sup that we have done freely needs justification.

4 Lower Bounds

We begin with some estimates on admissible controls and $u_\varepsilon(t,x,\omega)$. Recall the notation introduced in (3.5) and (3.6).

LEMMA 4.1 *Assume* (2.3) *and* (2.10)*. Then in the variational formula* (3.4)*, the supremum over $\mathcal{C}$ can be replaced with the supremum over the subset $\mathcal{C}^* \subset \mathcal{C}$ of controls that satisfy the following condition*: *for each $\delta > 0$ there is $C_\delta > 0$, which depends only on δ and the constants in* (2.3) *and* (2.10)*, such that*

$$\sup_{x,\omega} E^{Q_x^{\varepsilon,c}} |\xi_\varepsilon(t)| \leqslant C_\delta(t + \sqrt{\varepsilon t}) + 2\beta\delta. \tag{4.1}$$

In particular, for all $c \in \mathcal{C}^*$,

$$\sup_{x,\omega} E^{Q_x^{\varepsilon,c}} \left[\int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\|^{\beta'} ds \right] \leqslant C_\delta(t + \sqrt{\varepsilon t}) + 2\beta\delta. \tag{4.2}$$

PROOF: There is a trivial lower bound with $c \equiv 0$ that can be obtained from (2.3), (2.10), and (3.5): for every $\delta > 0$

$$\begin{aligned} u_\varepsilon(t,x,\omega) - f(x) &\geqslant E^{Q_x^{\varepsilon,0}}[f(x(t)) - f(x) - \xi_\varepsilon(t)] \\ &\geqslant -K_\delta E^{Q_x^{\varepsilon,0}}[\|x(t) - x\|] - c_4 t - \delta \\ &\geqslant -K_\delta \sqrt{\varepsilon t} - c_4 t - \delta. \end{aligned} \tag{4.3}$$

So we need to consider only those $c \in \mathcal{C}$ for which

$$E^{Q_x^{\varepsilon,c}}[f(x(t)) - f(x) - \xi_\varepsilon(t)] \geqslant -K_\delta\sqrt{\varepsilon t} - c_4 t - \delta.$$

In view of (2.10), such a c has to satisfy

$$E^{Q_x^{\varepsilon,c}}[K_\delta \|x(t) - x\| - \xi_\varepsilon(t)] \geqslant -K_\delta\sqrt{\varepsilon t} - c_4 t - 2\delta. \tag{4.4}$$

The following simple inequalities allow us to estimate $E^{Q_x^{\varepsilon,c}} \|x(t) - x\|$ in terms of $E^{Q_x^{\varepsilon,c}} \xi_\varepsilon(t)$:

$$E^{Q_x^{\varepsilon,c}} \|x(t) - x\| \leqslant E^{Q_x^{\varepsilon,c}} \left[\int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\| ds + \sqrt{\varepsilon t} \right], \tag{4.5}$$

$$\begin{aligned} &E^{Q_x^{\varepsilon,c}} \int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\| ds \\ &\qquad \leqslant t^{1/\beta} \left(E^{Q_x^{\varepsilon,c}} \int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\|^{\beta'} ds \right)^{1/\beta'}, \end{aligned} \tag{4.6}$$

$$E^{Q_x^{\varepsilon,c}} \int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\|^{\beta'} ds \leqslant c_1^{-1}\big(E^{Q_x^{\varepsilon,c}} \xi_\varepsilon(t) + c_2 t\big). \tag{4.7}$$

The last inequality follows from (2.3). Notice also that $\xi_\varepsilon(t) + c_2 t \geqslant 0$ and is nondecreasing for all t. Denoting $E^{Q_x^{\varepsilon,c}}(\xi_\varepsilon(t) + c_2 t)$ by $\Theta(t)$, we obtain from (4.4) that the set of controls $\mathcal{C}$ can be reduced to those for which

$$K_\delta\big[c_1^{-1/\beta'} t^{1/\beta} \Theta(t)^{1/\beta'} + \sqrt{\varepsilon t}\big] - \Theta(t) + K_\delta\sqrt{\varepsilon t} + (c_4 + c_2)t + 2\delta \geqslant 0. \tag{4.8}$$

The above relation is of the form $A\Theta^{1/\beta'} - \Theta + B \geqslant 0$. By Young's inequality this implies that

$$\frac{A^\beta}{\beta} + \frac{\Theta}{\beta'} - \Theta + B \geqslant 0.$$

Since $1/\beta + 1/\beta' = 1$, we conclude that

$$\Theta \leqslant A^\beta + \beta B.$$

This immediately gives (4.1). Inequality (4.2) follows from (4.1) and (4.7). □

LEMMA 4.2 *Assuming the bounds* (2.1) *and* (2.3)*, and* (2.10) *on* f*, there is a function* $c(t) \to 0$ *as* $t \to 0$ *such that for all* $0 < \varepsilon \leqslant 1$*,*

$$\sup_\omega |u_\varepsilon(t,x,\omega) - f(x)| \leqslant c(t). \tag{4.9}$$

PROOF: We start with the variational formula (3.4)

$$u_\varepsilon(t,x,\omega) - f(x) = \sup_{c\in\mathcal{C}^*} E^{Q_x^{\varepsilon,c}}[f(x(t)) - f(x) - \xi_\varepsilon(t)]$$

where $\mathcal{C}^*$ is as in Lemma 4.1. From (3.5), (2.10), and (4.5) we get for every $\delta > 0$

$$\begin{aligned} E^{Q_x^{\varepsilon,c}}&[f(x(t)) - f(x) - \xi_\varepsilon(t)] \\ &\leqslant E^{Q_x^{\varepsilon,c}}[K_\delta\|x(t) - x\| - \xi_\varepsilon(t)] + \delta \\ &\leqslant K_\delta\sqrt{\varepsilon t} + E^{Q_x^{\varepsilon,c}}\left[K_\delta\int_0^t \|c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\|ds - \xi_\varepsilon(t)\right] + \delta. \end{aligned}$$

Combining this with (4.1), we get an upper bound. Inequality (4.3) provides a lower bound. □

The starting point for lower bounds is the ergodic theorem. Let us begin with $(b,\varphi) \in \mathcal{E}$ and recall the definitions in (3.7) and (3.8):

$$m(b,\varphi) = \int b(\omega)\varphi(\omega)d\mathbb{P},$$
$$h(b,\varphi) = \int L(b(\omega),\omega)\varphi(\omega)d\mathbb{P}.$$

For the Markov process on Ω with generator

$$\mathcal{A}_b = \frac{1}{2}\Delta + \langle b(\omega), \nabla\rangle$$

$\varphi\, d\mathbb{P}$ is an ergodic invariant probability measure. By the ergodic theorem for almost all ω with respect to $\mathbb{P}$, the corresponding diffusion on Ω starting from ω, i.e., $P^{b,\omega}$, satisfies

$$\lim_{\varepsilon\to 0}\varepsilon\int_0^{t/\varepsilon} b(\omega(s))ds = t\int b(\omega)\varphi(\omega)d\mathbb{P} = m(b,\varphi)t,$$
$$\lim_{\varepsilon\to 0}\varepsilon\int_0^{t/\varepsilon} L(b(\omega(s)),\omega(s))ds = t\int L(b(\omega),\omega)\varphi(\omega)d\mathbb{P} = h(b,\varphi)t,$$

in $L^1(P^{b,\omega})$, uniformly in any finite t-interval. We can assume that this holds for every $\omega \in N$ where $\mathbb{P}(N) = 1$. In fact, by Egoroff's theorem, given any $\eta > 0$, there exists N_η with $\mathbb{P}(N_\eta) \geqslant 1 - \eta$, such that

$$\lim_{\varepsilon\to 0} \sup_{\omega\in N_\eta} \sup_{0\leqslant t\leqslant T} E^{P^{b,\omega}}\left(\left|\varepsilon \int_0^{t/\varepsilon} b(\omega(s))ds - m(b,\varphi)t\right|\right) = 0 \tag{4.10}$$

and

$$\lim_{\varepsilon\to 0} \sup_{\omega\in N_\eta} \sup_{0\leqslant t\leqslant T} E^{P^{b,\omega}}\left(\left|\varepsilon \int_0^{t/\varepsilon} L(b(\omega(s)),\omega(s))ds - h(b,\varphi)t\right|\right) = 0. \tag{4.11}$$

These properties can be expressed in terms of $Q_0^{b,\omega}$ as well:

$$\lim_{\varepsilon\to 0} \sup_{\omega\in N_\eta} \sup_{0\leqslant t\leqslant T} E^{Q_0^{b,\omega}}\left(\left|\varepsilon \int_0^{t/\varepsilon} b(\tau_{x(s)}\omega)ds - m(b,\varphi)t\right|\right) = 0, \tag{4.12}$$

$$\lim_{\varepsilon\to 0} \sup_{\omega\in N_\eta} \sup_{0\leqslant t\leqslant T} E^{Q_0^{b,\omega}}\left(\left|\varepsilon \int_0^{t/\varepsilon} L(b(\tau_{x(s)}\omega),\tau_{x(s)}\omega)ds - h(b,\varphi)t\right|\right) = 0. \tag{4.13}$$

LEMMA 4.3 *Assume* (H1) *and* (H6). *Let* $(b,\varphi) \in \mathcal{E}$ *and* $m = m(b,\varphi)$, $h = h(b,\varphi)$. *Then*

$$\liminf_{\varepsilon\to 0} \inf_{\omega\in N_\eta} \inf_{0\leqslant t\leqslant T} [u_\varepsilon(t,0,\omega) - f(mt) + ht] \geqslant 0. \tag{4.14}$$

PROOF: From the variational formula (3.4), we have for any $c \in \mathcal{C}$,

$$u_\varepsilon(t,0,\omega) \geqslant E^{Q_0^{c,\omega}}\left(f(\varepsilon x(\varepsilon^{-1}t)) - \varepsilon \int_0^{t/\varepsilon} L(c(s,x(s)),\tau_{x(s)}\omega)ds\right).$$

Choose $c(s,x) = c(x) = b(\tau_x\omega)$. Since f satisfies (2.10) and

$$x(t) = \int_0^t b(\tau_{x(s)}\omega)ds + \beta(t) = \int_0^t b(\omega(s))ds + \beta(t),$$

(4.14) follows from the ergodic theorem just as (4.10) and (4.11). □

LEMMA 4.4 *Assume* (H1), (H2), *and* (H6). *Let* $(b,\varphi) \in \mathcal{E}$, $m = m(b,\varphi)$, *and* $h = h(b,\varphi)$. *Then*

$$\lim_{r\to 0} \liminf_{\varepsilon\to 0} \inf_{\omega\in N_\eta} \inf_{0\leqslant t\leqslant T} \inf_{\|y\|\leqslant r} [u_\varepsilon(t,y,\omega) - f(mt) + ht] \geqslant 0. \tag{4.15}$$

PROOF: By Lemma 4.2, we can assume without loss of generality that $t \geqslant r$. We have a process $x(s)$ that starts from 0 having for its distribution $Q_0^c = Q_0^{b,\omega}$ with $c(s,x) = b(\tau_x\omega)$. It is enough to construct a process $y(s)$ with generator $\frac{1}{2}\Delta + \langle \hat{c}(s,\cdot), \nabla\rangle$ that starts from $\varepsilon^{-1}y$, which is close to $x(s)$ in the sense that

$$\varepsilon\|y(\varepsilon^{-1}t) - x(\varepsilon^{-1}t)\|$$

and

$$\varepsilon\left|\int_0^{t/\varepsilon} L(c(s,x(s)),\tau_{x(s)}\omega)ds - \int_0^{t/\varepsilon} L(\hat{c}(s,y(s)),\tau_{y(s)}\omega)ds\right|$$

are small. We will construct such a $y(\cdot)$ that is coupled to $x(\cdot)$.

Define

$$y(s) = \begin{cases} x(s) + y/\varepsilon - sy/\|y\| & \text{if } 0 \leqslant s \leqslant \varepsilon^{-1}\|y\| \\ x(s) & \text{if } s \geqslant \varepsilon^{-1}\|y\|. \end{cases}$$

Clearly,

$$\varepsilon\|y(\varepsilon^{-1}t) - x(\varepsilon^{-1}t)\| \leqslant \|y\| \leqslant r.$$

A simple calculation reveals

$$\hat{c}(s,z) = \begin{cases} c(s, z - y/\varepsilon + sy/\|y\|) - y/\|y\| & \text{if } 0 \leqslant s \leqslant \varepsilon^{-1}\|y\| \\ c(s,z) & \text{if } s \geqslant \varepsilon^{-1}\|y\|, \end{cases}$$

and $\|\hat{c}(\cdot,\cdot)\|_\infty \leqslant \|c(\cdot,\cdot)\|_\infty + 1$. We have

$$\begin{aligned} &\varepsilon\left|\int_0^{t/\varepsilon} L(c(s,x(s)),\tau_{x(s)}\omega)ds - \int_0^{t/\varepsilon} L(\hat{c}(s,y(s)),\tau_{y(s)}\omega)ds\right| \\ &\quad = \varepsilon\left|\int_0^{\|y\|/\varepsilon} L(c(s,x(s)),\tau_{x(s)}\omega)ds - \int_0^{\|y\|/\varepsilon} L(\hat{c}(s,y(s)),\tau_{y(s)}\omega)ds\right| \\ &\quad \leqslant 2\|y\| \sup_{|p|\leqslant\|c(\cdot,\cdot)\|_\infty+1} \sup_\omega |L(p,\omega)| \end{aligned}$$

The statement of the lemma follows from the above estimates and (4.14). □

We now turn to the proof of Theorem 2.1.

PROOF: If we want to show that $u_\varepsilon(t,y,\omega) \to u(t,y)$ in probability, we could translate by $\tau_{\varepsilon^{-1}y}$ and the initial condition would be $f_y(x) = f(x+y)$. Since $\mathbb{P}$ is invariant, statistically nothing would have changed. However, it is not obvious that the convergence is almost sure. Nor is it obvious that the convergence is locally uniform in x.

Let us remark that our proof of (4.14) depended only on the constant K_δ in (2.10). Also note that the constant K_δ corresponding to f also works for any translate $f^y(x) = f(x+y)$. As a result, if we decide to start from 0, but consider for some $\ell \geqslant 1$ the family $\{f^y(x) = f(x+y) : \|y\| \leqslant \ell\}$ of initial functions and their corresponding solutions $u_\varepsilon^y(t,x,\omega)$, then as we saw in (4.14), for any given $\eta > 0$, there is a set N_η, with $\mathbb{P}(N_\eta) \geqslant 1-\eta$, such that

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} \sup_{\|y\|\leqslant\ell} \sup_{\omega\in N_\eta} [f(y+mt) - th - u_\varepsilon^y(t,0,\omega)]^+ = 0.$$

By the ergodic theorem, there is a set N with $\mathbb{P}(N) = 1$ and for $\omega \in N$,

$$\lim_{\varepsilon\to 0} \frac{|\{x : \tau_x\omega \in N_\eta, \|x\| \leqslant \ell\varepsilon^{-1}\}|}{|\{x : \|x\| \leqslant \ell\varepsilon^{-1}\}|} = \mathbb{P}(N_\eta) \geqslant 1-\eta.$$

Let $\omega \in N$. Then for $\varepsilon \leqslant \varepsilon_0(\eta)$,

$$\big|\{x : \tau_x\omega \in N_\eta, \|x\| \leqslant \ell\varepsilon^{-1}\}\big| \geqslant (1-2\eta)\big|\{x : \|x\| \leqslant \ell\varepsilon^{-1}\}\big|.$$

In particular, if $\delta(\eta) = \ell(3\eta)^{1/d}$, then every x satisfying $\|x\| \leqslant \varepsilon^{-1}\ell$ is within a distance $\delta(\eta)\varepsilon^{-1}$ of an x' such that $\tau_{x'}\omega \in N_\eta$. Now from (4.15) and the relation (3.2) it follows easily that

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} \sup_{\|x\|\leqslant \ell} [f(x+mt) - ht - u_\varepsilon(t,x,\omega)]^+ = 0$$

for $\omega \in \bigcup_{\eta>0} N_\eta$. If we denote by $\mathbf{A} \in \mathbb{R}^d \times \mathbb{R}$ the range of $(m(b,\varphi), h(b,\varphi))$ as (b,φ) varies over $\mathcal{E}$, then it is routine to show that with

$$u(t,x) = \sup_{(m,h)\in\mathbf{A}} [f(x+mt) - ht],$$

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} \sup_{\|x\|\leqslant \ell} [u(t,x) - u_\varepsilon(t,x,\omega)]^+ = 0$$

$\mathbb{P}$ a.s., which proves Theorem 2.1. □

5 Convex Analysis

We start with a simple estimate. For any $\varepsilon > 0$, there exists a constant C_ε such that

$$\|m\| \leqslant C_\varepsilon + \varepsilon h$$

for all $(m,h) \in \mathbf{A}$, which is a consequence of (2.3). Therefore, for any $\theta \in \mathbb{R}^d$,

$$\overline{H}(\theta) = \sup_{(m,h)\in\mathbf{A}} [\langle\theta, m\rangle - h] \tag{5.1}$$

is finite. The aim of this section is to show that, for each $\theta \in \mathbb{R}^d$, there exists a "function" $\psi(\omega) = \psi_\theta(\omega)$ such that

$$\frac{1}{2}\Delta\psi(\omega) + H(\nabla\psi(\omega), \omega) \leqslant \overline{H}(\theta)$$

for almost all ω. The function ψ does not really exist, but $\mathbf{W} = \nabla\psi$ exists as a function in some $L^p(\Omega)$ with values in $\mathbb{R}^d$. Moreover, $\nabla \times \mathbf{W} = 0$ and

$$\frac{1}{2}\nabla\cdot\mathbf{W} + H(\mathbf{W}, \omega) \leqslant \overline{H}(\theta)$$

in the sense of distributions.

We now represent $\mathcal{E}$ as a limit of an increasing sequence of compact sets. We can assume without loss of generality that $L^p(\Omega,\mathbb{P})$ are all separable spaces for $1 \leqslant p < \infty$. Let L^1_+ be the convex subset of $\varphi \in L^1(\Omega,\mathbb{P})$ consisting of $\varphi(\cdot)$ that satisfy $\varphi(\omega) \geqslant 0$ a.e. $\mathbb{P}$ and $\int_\Omega \varphi(\omega)\,d\mathbb{P} = 1$. Let $\{\varphi_j(\omega)\}$ be a countable dense subset in L^1_+ in the strong L^1 topology. We will also assume that each $\varphi_j(\omega)$ is smooth in the sense that $\varphi_j(\tau_x\omega) = \varphi_j(x,\omega)$ is uniformly bounded above and below (away from 0) on $\mathbb{R}^d$ and has two bounded continuous derivatives in x. The bounds will be independent of ω but could and in fact will depend on j. We denote by $\mathbf{D}_k$ the convex hull of $\{\varphi_j : 1 \leqslant j \leqslant k\}$. In particular, for each k, $\mathbf{D}_k$ is a compact subset of $L^1_+ \subset L^1(\Omega,\mathbb{P})$. Each $\varphi \in \mathbf{D}_k$ has uniform bounds

$$\varphi(\omega) \geqslant c_k, \quad \|\nabla\varphi\| \leqslant C_k, \quad \|\nabla^2\varphi\| \leqslant C_k.$$

Now we consider maps $b : \Omega \to \mathbb{R}^d$. Set

$$\begin{aligned}
\mathbf{B}_r &= \{b : \sup_\omega \|b(\omega)\| \leqslant r\},\\
\mathcal{E}_{r,k} &= \big\{(b,\varphi) : b \in \mathbf{B}_r,\ \varphi \in \mathbf{D}_k,\ \tfrac{1}{2}\Delta\varphi = \nabla\cdot(b\,\varphi)\big\},\\
\mathcal{E}_k &= \bigcup_r \mathcal{E}_{r,k},\\
\mathcal{E} &= \bigcup_k \mathcal{E}_k.
\end{aligned}$$

In the definition of $\mathcal{E}_{r,k}$ the equation $\frac{1}{2}\Delta\varphi = \nabla\cdot(b\,\varphi)$ is interpreted in the weak sense. This is sufficient to show that the Markov process $P^{b,\omega}$, which is well-defined, has $\varphi\, d\mathbb{P}$ as an ergodic invariant measure. Let us note that while each $\mathbf{D}_k$ is strongly compact, each $\mathbf{B}_r$ is weakly compact.

For each $\theta \in \mathbb{R}^d$ we have

$$\overline{H}(\theta) \geqslant \sup_{(b,\varphi)\in\mathcal{E}_{r,k}} \left[\int [\langle\theta, b(\omega)\rangle - L(b(\omega),\omega)]\,\varphi(\omega) d\mathbb{P}\right] \tag{5.2}$$

$$= \sup_{\varphi\in\mathbf{D}_k} \sup_{b\in\mathbf{B}_r} \inf_{u\in\mathcal{U}} \left[\int [\langle\theta, b(\omega)\rangle - L(b(\omega),\omega)]\varphi(\omega)\, d\mathbb{P} + \int (\mathcal{A}_b u)\varphi(\omega) d\mathbb{P}\right] \tag{5.3}$$

$$= \sup_{\varphi\in\mathbf{D}_k} \inf_{u\in\mathcal{U}} \sup_{b\in\mathbf{B}_r} \left[\int [\langle\theta, b(\omega)\rangle - L(b(\omega),\omega)]\varphi(\omega)\, d\mathbb{P} + \int (\mathcal{A}_b u)\,\varphi(\omega)\, d\mathbb{P}\right] \tag{5.4}$$

$$= \sup_{\varphi\in\mathbf{D}_k} \inf_{u\in\mathcal{U}} \left[\int \left(\frac{1}{2}\Delta u + H_r(\theta+\nabla u,\omega)\right)\varphi\, d\mathbb{P}\right] \tag{5.5}$$

$$= \inf_{u\in\mathcal{U}} \sup_{\varphi\in\mathbf{D}_k} \left[\int \left(\frac{1}{2}\Delta u + H_r(\theta+\nabla u,\omega)\right)\varphi\, d\mathbb{P}\right] \tag{5.6}$$

Since $\mathcal{E} \supset \mathcal{E}_{r,k}$, (5.2) is obvious. On the other hand, in (5.2), b and φ have to be related by $\frac{1}{2}\Delta\varphi = \nabla\cdot(b\,\varphi)$. This is taken care of by the term $\mathcal{A}_b u = \frac{1}{2}\Delta u + \langle b, \nabla u\rangle$ and

$$\inf_{u\in\mathcal{U}} \int (\mathcal{A}_b u)\varphi\, d\mathbb{P} = -\infty$$

unless $\frac{1}{2}\Delta\varphi = \nabla\cdot(b\,\varphi)$ in the weak sense, in which case it is 0. For this we can take any reasonable linear space $\mathcal{U}$ of test functions. This establishes (5.3). The functional on the right is clearly a concave function of $b \in \mathbf{B}_r$ and is upper semicontinuous in the weak topology in which $\mathbf{B}_r$ is compact. It is linear and continuous in $u \in \mathcal{U}$. We can apply the minimax theorem (see [18]) and interchange the order of $\inf_u$ and $\sup_b$ for each fixed φ. We thus arrive at (5.4).

Now, the expression inside the integral is a local expression in $b(\omega)$ and the supremum can, therefore, be taken inside the integral. Setting

$$H_r(\xi, \omega) = \sup_{\|q\| \leqslant r} [\langle \xi, q \rangle - L(q, \omega)],$$

we get (5.5). The expression in (5.5) is convex and continuous in $u \in \mathcal{U}$ and linear and continuous in $\varphi \in \mathbf{D}_k$. Moreover, $\mathbf{D}_k$ is compact. We can use the minimax theorem once more to arrive at (5.6).

We have therefore proved the following:

LEMMA 5.1 *Let $\theta \in \mathbb{R}^d$ be given. For each integer $k \geqslant 1$ there exists a function $u_k(\omega) \in \mathcal{U}$ such that*

$$\sup_{\varphi \in \mathbf{D}_k} \int_\Omega \left(\frac{1}{2} \Delta u_k + H_k(\theta + \nabla u_k, \omega) \right) \varphi \, d\mathbb{P} \leqslant \overline{H}(\theta) + \frac{1}{k}.$$

The next step is to obtain some sort of a "weak limit" u from $\{u_k\}$ such that

$$\sup_k \sup_{\varphi \in \mathbf{D}_k} \int_\Omega \left(\frac{1}{2} \Delta u + H(\theta + \nabla u, \omega) \right) \varphi \, d\mathbb{P} \leqslant \overline{H}(\theta).$$

The function u may not exist, but we will show that $\nabla u = v$ exists.

Let $v_k = \nabla u_k$. Notice that the constant function $1 \in \mathbf{D}_k$ for every k, so we can choose $\varphi \equiv 1$. Since $\int \Delta u \, d\mathbb{P} = 0$ for any $u \in \mathcal{U}$, we get the bound

$$\int H_k(\theta + v_k(\omega), \omega) d\mathbb{P} \leqslant \overline{H}(\theta) + \frac{1}{k}. \tag{5.7}$$

THEOREM 5.2 *The sequence $\{v_k\}$ is uniformly integrable in $L^1(\Omega, \mathbb{P})$, and any weak limit v of v_k satisfies*

$$\int v \, d\mathbb{P} = 0, \quad \nabla \times v = 0,$$

$$\sup_{\varphi \in \bigcup_k \mathbf{D}_k} \int_\Omega \left(\frac{1}{2} \nabla \cdot v + H(\theta + v, \omega) \right) \varphi \, d\mathbb{P} \leqslant \overline{H}(\theta),$$

and hence also satisfies

$$\frac{1}{2} (\nabla \cdot v)(\omega) + H(\theta + v(\omega), \omega) \leqslant \overline{H}(\theta) \tag{5.8}$$

as a distribution.

PROOF: Once we establish uniform integrability, the rest is routine. If v is a weak limit point of $\{v_k\}$, convexity of H_k guarantees that

$$\sup_{\varphi\in\mathbf{D}_k}\int_\Omega\left(\frac{1}{2}\nabla\cdot v+H_k(\theta+v,\omega)\right)\varphi\,d\mathbb{P}\leqslant\overline{H}(\theta).$$

Since the $\mathbf{D}_k$ are increasing and $H_k\uparrow H$, we can replace H_k by H. Since $\bigcup_k\mathbf{D}_k$ is dense in L^1_+, the uniform bound on the integrals turns into an almost sure bound on the integrand.

Let us now establish the uniform integrability of $\{v_k\}$. We can get a lower bound on H_k from (2.3),

$$H_k(\xi,\omega)\geqslant\sup_{\|y\|\leqslant k}[\langle y,\xi\rangle-c_3-c_4\|y\|^{\alpha'}]=\sup_{0\leqslant y\leqslant k}[y\|\xi\|-c_3-c_4y^{\alpha'}].$$

Let c be a small positive number. If $\|\xi\|\leqslant(k/c)^{\alpha'-1}$, we pick $y=c\|\xi\|^{1/(\alpha'-1)}$ and get the lower bound

$$H_k(\xi,\omega)\geqslant c\|\xi\|^\alpha-c_3-c_4c^{\alpha'}\|\xi\|^\alpha\geqslant c_5\|\xi\|^\alpha-c_3,$$

where $1/\alpha'+1/\alpha=1$ and $c_5>0$ provided that c is small enough. If $\|\xi\|>(k/c)^{\alpha'-1}$, then we pick $y=k$ and obtain for all sufficiently small c

$$H_k(\xi,\omega)\geqslant k\|\xi\|-c_3-c_4k^{\alpha'}\geqslant k\|\xi\|-c_3-c_4kc^{\alpha'-1}\|\xi\|\geqslant c_6k\|\xi\|-c_3.$$

From the above lower bounds we conclude that

$$\begin{aligned}\int_{\|\xi\|\geqslant A}\|\xi\|d\mathbb{P}&=\int_{\substack{\|\xi\|\geqslant A\\ \|\xi\|\leqslant(k/c)^{\alpha'-1}}}\|\xi\|d\mathbb{P}+\int_{\substack{\|\xi\|\geqslant A\\ \|\xi\|\geqslant(k/c)^{\alpha'-1}}}\|\xi\|d\mathbb{P}\\&\leqslant\int_{\substack{\|\xi\|\geqslant A\\ \|\xi\|\leqslant(k/c)^{\alpha'-1}}}\frac{\|\xi\|^\alpha}{\|\xi\|^{\alpha-1}}d\mathbb{P}+\int_{\|\xi\|\geqslant(k/c)^{\alpha'-1}}\|\xi\|d\mathbb{P}\\&\leqslant\frac{1}{c_5A^{\alpha-1}}\int[H_k(\xi,\omega)+c_3]d\mathbb{P}+\frac{1}{kc_6}\int[H_k(\xi,\omega)+c_3]d\mathbb{P}.\end{aligned}$$

Setting $\xi=\theta+v_k$ establishes the uniform integrability of $\{v_k\}$. □

6 Upper Bounds: Preliminaries

Now we use the function v constructed in (5.8) to provide us with an upper bound with the help of the maximum principle. The idea is, roughly speaking, to use $v(x,\omega)=v(\tau_x\omega)$ as a gradient to obtain $V(x,\omega)$ normalized by the condition $V(0,\omega)=0$ for $\mathbb{P}$ a.e. ω. Observe that, since $v\in L^\alpha(\Omega)$ is irrotational in the weak sense, it has a potential $V(\,\cdot\,,\omega)\in W^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^d)$ for a.e. ω. In fact,

$$V(x,\omega)=\int_{0\to x}\langle v,dz(s)\rangle,\tag{6.1}$$

where $z(s)$ is a path connecting 0 to x. It is now easy to estimate

$$\|V(x,\cdot) - V(y,\cdot)\|_\alpha \leqslant C\|x-y\|.$$

By our choice of the normalization,

$$\|V(x,\cdot)\|_\alpha \leqslant C\|x\|. \tag{6.2}$$

If v were nice, the ergodic theorem would tell us that $V(x,\omega) = o(\|x\|)$ as $\|x\| \to \infty$ a.s.

The function $\widehat{V}_\varepsilon(t,x,\omega) = \langle\theta, x\rangle + t\overline{H}(\theta) + \varepsilon V(x/\varepsilon,\omega)$ would be a supersolution of (1.1). Its initial value $\widehat{V}_\varepsilon(0,x,\omega) = \langle\theta,x\rangle + \varepsilon V(x/\varepsilon,\omega)$ would differ from the affine function $\langle\theta,x\rangle$ by a vanishing term $\varepsilon V(x/\varepsilon,\omega) = o(1)$ as $\varepsilon \to 0$. Moreover, setting $x = 0$ and using our normalization $V(0,\omega) = 0$, we would get by comparison that with probability 1

$$\limsup_{\varepsilon\to 0} u_\varepsilon(t,0,\omega) \leqslant t\overline{H}(\theta) = u(t,0),$$

establishing the upper bound for the affine initial data at $x = 0$. We shall see in Corollary 6.8 at the end of this section that in this case the homogenization takes place under weak hypotheses (H1) and (H2) without assuming the regularity of v. Additional conditions will be needed to strengthen this result to the locally uniform bound and general initial data.

At first, we need to understand the nature of the solution v. It is clear from the lower bound on H that $v \in L^\alpha(\Omega,\mathbb{P})$ where $\alpha > 1$ is as in (2.1). Then for almost all ω with respect to $\mathbb{P}$, $v(x,\omega) = v(\tau_x\omega) \in L^\alpha_{\text{loc}}(\mathbb{R}^d)$ and satisfies as a distribution on $\mathbb{R}^d$,

$$\frac{1}{2}(\nabla\cdot v)(\tau_x\omega) + H(\theta + v(\tau_x\omega), \tau_x\omega) \leqslant \overline{H}(\theta). \tag{6.3}$$

LEMMA 6.1 *For any $r > 0$, there is a constant $C(r)$ independent of ω such that*

$$\sup_x \int_{\|y-x\|\leqslant r} \|v(y,\omega)\|^\alpha\, dy = C(r) < \infty$$

for almost all ω with respect to $\mathbb{P}$.

PROOF: Since the proof is carried out for a fixed ω, we shall drop ω from the notation. Let g be any compactly supported nonnegative test function such that $\int g(x)dx = 1$. From (6.3) and the lower bound in (2.1) we have

$$\int (\nabla\cdot v(y) + c_1\|\theta + v(y)\|^\alpha - c_2)g(y)dy \leqslant \overline{H}(\theta).$$

From this inequality we get

$$\int \|v(y)\|^\alpha g(y)dy \leqslant C_1 \int \langle(\nabla g)(y), v(y)\rangle dy + C_2$$

with constants C_1 and C_2 independent of g and ω. Let us take a g supported on $B(0, 2r)$ with

$$\int \left(\frac{\|\nabla g\|}{g}\right)^{\alpha'} g(y)dy = C_3(r) < \infty \quad \text{and} \quad g(y) \geqslant c(r) > 0 \text{ on } B(0,r).$$

Then an application of Young's inequality yields

$$\begin{aligned}\int \|v(y)\|^\alpha g(y)dy &\leqslant C_1 \int \left\langle \frac{(\nabla g)(y)}{g(y)}, v(y)\right\rangle g(y)dy + C_2\\ &\leqslant \frac{1}{2}\int \|v(y)\|^\alpha g(y)dy + C_4(r),\end{aligned}$$

providing us with a uniform estimate

$$\int_{B(0,r)} \|v(y)\|^\alpha dy \leqslant C(r).$$

Since replacing $g(y)$ by $g(x-y)$ does not change anything, the estimate is uniform in x (which is the same as being uniform in ω). □

We now try to improve the $L^\alpha(\Omega, \mathbb{P})$ bound on v, replacing it with a convolution $\hat{v}_\delta \in L^\infty(\Omega, \mathbb{P})$. The proof of Lemma 6.2 below is straightforward and omitted.

LEMMA 6.2 *Let us define the convolution*

$$v_\delta(\omega) = \int v(\tau_{\delta y}\,\omega)\rho(y)dy \tag{6.4}$$

where $\rho(y)$ is a mollifier supported on $\|y\| \leqslant 1$. Then v_δ satisfies

$$\|v_\delta\|_\infty \leqslant C_\delta$$

for some constant C_δ. If we now define

$$V_\delta(x) = \int V(x+\delta y)\rho(y)dy \tag{6.5}$$

as the corresponding convolution of V, then $\nabla V_\delta = v_\delta$ and

$$\|V_\delta(0,\omega)\|_\infty \leqslant C_\delta$$

with a possibly different C_δ. Moreover, by the ergodic theorem for any $\delta > 0$ and $\ell < \infty$,

$$\lim_{\varepsilon\to 0} \varepsilon \sup_{|x|\leqslant \ell} |V_\delta(\varepsilon^{-1}x, \omega)| = 0 \quad a.e.\ \mathbb{P}.$$

Under (H3) one can use the mollified functions V_δ for comparison.

LEMMA 6.3 *Assume the strong hypothesis* (H3). *Then for each $\delta > 0$, there exists a smooth $\hat{v}_\delta$ that is uniformly bounded in ω, satisfies $\nabla \times \hat{v}_\delta = 0$, $\mathbb{E}[\hat{v}_\delta] = 0$, and*

$$\frac{1}{2}\nabla \cdot \hat{v}_\delta + H(\theta + \hat{v}_\delta, \omega) \leqslant \overline{H}(\theta) + \hat{\nu}(\delta),$$

where $\hat{\nu}(\delta) \to 0$ as $\delta \to 0$.

PROOF: We start with the convolution v_δ as in (6.4). Clearly,

$$\frac{1}{2}\nabla\cdot v_\delta + \int H(\theta + v(\tau_{\delta y}\omega), \tau_{\delta y}\omega)\rho(y)dy \leqslant \overline{H}(\theta),$$

and (2.7) allows us to conclude that

$$\frac{1}{2}\nabla\cdot v_\delta + (1+\nu(\delta))\int H((1+\nu(\delta))^{-1}(\theta + v(\tau_{\delta y}\omega)), \omega)\rho(y)dy \leqslant \overline{H}(\theta) + C\nu(\delta).$$

The convexity of H can now be used to infer that $\hat{v}_\delta = v_\delta/(1+\nu(\delta))$ satisfies

$$\frac{1}{2}\nabla\cdot \hat{v}_\delta + H((1+\nu(\delta))^{-1}\theta + \hat{v}_\delta, \omega) \leqslant (1+\nu(\delta))^{-1}(\overline{H}(\theta) + C\nu(\delta)).$$

We can replace θ by $(1+\nu(\delta))\theta$ and use the continuity of $\overline{H}$, which is a convex function of θ. □

Next we prove a technical lemma that will allow us to use v_δ and V_δ and avoid dealing with weak supersolutions in the comparison arguments that follow.

LEMMA 6.4 *Assume* (H1) *and* (H2). *Let $\theta \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$ be such that there exists a mean zero irrotational vector field $v \in L^\alpha(\Omega)$, which is a weak solution of*

$$\frac{1}{2}(\nabla\cdot v)(\omega) + H(\theta + v(\omega), \omega) \leqslant \lambda.$$

Denote by V its integral normalized by the condition $V(0,\omega) = 0$ ω-a.s. Let $c \in \mathcal{C}$,

$$x(t) = x + \int_0^t c(s, x(s))ds + \beta(t), \quad \xi(t) = \int_0^t L(c(s,x(s)), \tau_{x(s)}\omega)ds.$$

Then there exists $r_c(\delta) \to 0$ as $\delta \to 0$, which depends only on the bound of $c \in \mathcal{C}$, such that

$$\langle\theta, x(t) - x\rangle - t\lambda - \xi(t) \leqslant V_\delta(x,\omega) - V_\delta(x(t),\omega) + \int_0^t \langle\theta + v_\delta(\tau_{x(s)}\omega), d\beta(s)\rangle + r_c(\delta)t.$$

PROOF: Applying Itô's formula to $V_\delta(x,\omega) + \langle\theta, x\rangle$, we get

$$\begin{aligned} &V_\delta(x(t),\omega) - V_\delta(x,\omega) + \langle\theta, x(t)-x\rangle \\ &= \int_0^t \langle\theta + v_\delta(\tau_{x(s)}\omega), c(s,x(s))\rangle + \frac{1}{2}\nabla\cdot v_\delta(\tau_{x(s)}\omega)ds \\ &\quad + \int_0^t \langle\theta + v_\delta(\tau_{x(s)}\omega), d\beta(s)\rangle. \end{aligned} \tag{6.6}$$

From the relationship between H and L it is obvious that for all c and p

$$\langle p, c\rangle \leqslant H(p,\omega) + L(c,\omega).$$

Thus, for every $x \in \mathbb{R}^d$,

$$
\begin{aligned}
&\langle \theta + v_\delta(\tau_x\omega), c(s,x)\rangle \\
&\quad = \int_{\mathbb{R}^d} \langle \theta + v(\tau_y\omega), c(s,x)\rangle \rho_\delta(x-y)dy \\
&\quad \leqslant \int_{\mathbb{R}^d} (L(c(s,x),\tau_y\omega) + H(\theta + v(\tau_y\omega),\tau_y\omega))\rho_\delta(x-y)dy.
\end{aligned}
\tag{6.7}
$$

Substituting this into (6.6) and using (6.3), we obtain

$$
\begin{aligned}
&V_\delta(x(t),\omega) - V_\delta(x,\omega) + \langle \theta, x(t)-x\rangle \\
&\quad \leqslant t\lambda + \int_0^t \int_{\mathbb{R}^d} L(c(s,x(s)),\tau_y\omega)\rho_\delta(x(s)-y)dy\,ds \\
&\qquad + \int_0^t \langle \theta + v_\delta(\tau_{x(s)}\omega), d\beta(s)\rangle.
\end{aligned}
$$

Rearranging the terms and subtracting $\xi(t)$ from both sides, we arrive at the inequality

$$
\begin{aligned}
&\langle \theta, x(t)-x\rangle - t\lambda - \xi(t) \\
&\quad \leqslant V_\delta(x,\omega) - V_\delta(x(t),\omega) + \int_0^t \langle \theta + v_\delta(\tau_{x(s)}\omega), d\beta(s)\rangle \\
&\qquad + \int_0^t \int_{\mathbb{R}^d} \big(L(c(s,x(s)),\tau_y\omega) - L(c(s,x(s)),\tau_{x(s)}\omega)\big)\rho_\delta(x(s)-y)dy\,ds.
\end{aligned}
$$

Since the drift c is bounded and L satisfies (H2), the last term is bounded in absolute value by $r_c(\delta)t$ where

$$
r_c(\delta) = \sup_{\|q\|\leqslant\|c(\cdot,\cdot)\|_\infty} \sup_\omega \sup_{\|y\|\leqslant\delta} |L(q,\tau_y\omega) - L(q,\omega)|
$$

tends to 0 as $\delta \to 0$. $\square$

As the first application, we obtain an upper bound at $x = 0$ for affine initial data under weak assumptions (H1) and (H2). We will need the following elementary fact, which we state as a lemma:

LEMMA 6.5 *Let $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $\frac{1}{2}\Delta g \leqslant C$ in the sense of distributions. Then there exists g_*, a lower semicontinuous modification of g, such that for every $\delta > 0$*

$$
g_*(x) \geqslant g_\delta(x) - \frac{C}{d}\delta^2 \quad \textit{for all } x \in \mathbb{R}^d.
$$

PROOF: It is a standard fact for $C = 0$. Subtracting a quadratic function from g leads to the above result. □

Remark 6.6. Below when we refer to V or V_ε, we shall always mean a lower semicontinuous modification given by Lemma 6.5 and will not use the subscript $*$.

THEOREM 6.7 *Assume* (H1) *and* (H2). *Let* $\theta \in \mathbb{R}^d$ *and* $\lambda \in \mathbb{R}$ *be as in Lemma* 6.4. *Let* $u_\varepsilon(t, x, \omega)$ *be the solution of* (1.1) *with the initial condition* $f(x) = \langle \theta, x \rangle$. *Then with probability* 1

$$\limsup_{\varepsilon \to 0} u_\varepsilon(t, 0, \omega) \leqslant t\lambda.$$

PROOF: Let V be as in Lemma 6.4 and V_δ be given by (6.5). Set $V_{\delta,\varepsilon}(x, \omega) = \varepsilon V_\delta(x/\varepsilon, \omega)$. Using Lemma 6.4 with $x = 0$, rescaling, and taking the expectation with respect to $Q_0^{\varepsilon,c}$, we get

$$E^{Q_0^{\varepsilon,c}}(\langle \theta, x(t) \rangle - \xi_\varepsilon(t)) \leqslant t\lambda - E^{Q_0^{\varepsilon,c}} V_{\delta,\varepsilon}(x(t), \omega) + V_{\delta,\varepsilon}(0, \omega) + r_c(\delta)t \tag{6.8}$$

where $r_c(\delta) \to 0$ as $\delta \to 0$ for each $c \in \mathcal{C}$.

We claim that with $\mathbb{P}$ probability 1

$$\lim_{\delta \to 0} E^{Q_0^{\varepsilon,c}} |V_{\delta,\varepsilon}(x(t), \omega) - V_\varepsilon(x(t), \omega)| = 0. \tag{6.9}$$

Let $p_{\varepsilon,c}(t, y)$ be the probability density of $x(t)$ under $Q_0^{\varepsilon,c}$. Then

$$E^{Q_0^{\varepsilon,c}} |V_{\delta,\varepsilon}(x(t), \omega) - V_\varepsilon(x(t), \omega)| = \int_{\mathbb{R}^d} |V_{\delta,\varepsilon}(y, \omega) - V_\varepsilon(y, \omega)| p_{\varepsilon,c}(t, y) dy.$$

Since c is bounded and $\varepsilon > 0$, $p_{\varepsilon,c}(t, x, y)$ has sufficient regularity and decay to imply (6.9) in view of (6.2). This allows us to pass to the limit as $\delta \to 0$ in (6.8) for each $\varepsilon > 0$ and obtain

$$E^{Q_0^{\varepsilon,c}}(\langle \theta, x(t) \rangle - \xi_\varepsilon(t)) \leqslant t\lambda - E^{Q_0^{\varepsilon,c}} V_\varepsilon(x(t), \omega) + V_\varepsilon(0, \omega).$$

Recalling that $V_\varepsilon(0, \omega) = 0$ a.s. and applying the lower bound of Lemma 6.5 to V_ε, we see that for every $\delta > 0$

$$E^{Q_0^{\varepsilon,c}}(\langle \theta, x(t) \rangle - \xi_\varepsilon(t)) \leqslant t\lambda - E^{Q_0^{\varepsilon,c}} V_{\delta,\varepsilon}(x(t), \omega) + \frac{C}{d}\, \varepsilon\delta^2.$$

From the variational formula and the above inequality, we get

$$u_\varepsilon(t, 0, \omega) \leqslant t\lambda - \inf_{c \in \mathcal{C}^*} E^{Q_0^{\varepsilon,c}} V_{\delta,\varepsilon}(x(t), \omega) + \frac{C}{d}\, \varepsilon\delta^2.$$

By Lemma 6.2, $V_\delta(x, \omega)$ has a bounded gradient. Then by the ergodic theorem $V_\delta(x, \omega)$ is sublinear: for every $\eta > 0$ there is a constant $C_{\eta,\delta}$ such that $|V_\delta(x, \omega)| \leqslant \eta \|x\| + C_{\eta,\delta}$ for all x and a.e. ω. After the rescaling we get

$$|V_{\delta,\varepsilon}(x, \omega)| \leqslant \eta \|x\| + \varepsilon C_{\eta,\delta}.$$

This inequality and Lemma 4.1 (see also (7.1) below) imply the desired statement when we let $\varepsilon \to 0$, then $\delta \to 0$, and finally $\eta \to 0$. □

By (6.3) we can choose $\lambda = \overline{H}(\theta)$. Using Theorem 6.7 and the already proved general lower bound, we obtain the homogenization result for affine initial data at $x = 0$.

COROLLARY 6.8 *Assume* (H1) *and* (H2). *Let* $u_\varepsilon(t, x, \omega)$ *be the solution of* (1.1) *with the initial condition* $f(x) = \langle \theta, x \rangle$. *Then with probability* 1

$$\limsup_{\varepsilon \to 0} u_\varepsilon(t, 0, \omega) = t\overline{H}(\theta).$$

The following obvious corollary will be used in the next section:

COROLLARY 6.9 *Let* $\theta \in \mathbb{R}^d$ *and* $\lambda \in \mathbb{R}$ *admit a* v *as in Lemma* 6.4. *Then* $\overline{H}(\theta) \leqslant \lambda$.

7 Proof of Theorem 2.2

We are now ready to proceed with the proof of Theorem 2.2. Let us recall the formula

$$u_\varepsilon(t, x, \omega) = \sup_{c \in \mathcal{C}} E^{Q_x^{\varepsilon,c}}[f(x(t)) - \xi_\varepsilon(t)],$$

where

$$\xi_\varepsilon(t) = \int_0^t L\big((c(\varepsilon^{-1}s, \varepsilon^{-1}x(s)), \tau_{\varepsilon^{-1}x(s)}\omega\big)ds$$

and $Q_x^{\varepsilon,c}$ is the diffusion process on $C([0, T]; \mathbb{R}^d)$ corresponding to the generator

$$\frac{\varepsilon}{2}\Delta + c(\varepsilon^{-1}s, \varepsilon^{-1}x) \cdot \nabla$$

starting at time 0 from $x \in \mathbb{R}^d$ so that almost surely with respect to $Q_x^{\varepsilon,c}$

$$x_\varepsilon(t) = x(t) = x + \int_0^t c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))ds + \sqrt{\varepsilon}\beta(t),$$

$\beta(\cdot)$ being the standard Brownian motion on $\mathbb{R}^d$.

The following steps require only assumptions (H1), (H2), and (H6).

Step 1. As we saw in Lemma 4.1, the supremum in the variational formula (3.4) can be restricted to the set $\mathcal{C}^*$ that consists of all controls $c \in \mathcal{C}$ that satisfy condition (4.1). Observe that $\|x_\varepsilon(t)\|$ and therefore $f(x_\varepsilon(t))$ is uniformly integrable with respect to $\{Q_x^{\varepsilon,c} : \|x\| \leqslant \ell,\ 0 \leqslant t \leqslant T,\ 0 < \varepsilon \leqslant 1,\ c \in \mathcal{C}^*\}$, and

$$\sup_{\|x\| \leqslant \ell} \sup_{0 \leqslant t \leqslant T} \sup_{c \in \mathcal{C}^*} \sup_{0 < \varepsilon \leqslant 1} E^{Q_x^{\varepsilon,c}}[\|x_\varepsilon(t)\| + |\xi_\varepsilon(t)|] < \infty. \tag{7.1}$$

This is a simple consequence of Lemma 4.1 and inequalities (4.5) and (4.6).

Step 2. For any $\delta > 0$, we can find M_δ such that

$$\begin{aligned} &u_\varepsilon(t, x, \omega) - u(t, x) \\ &\quad = \sup_{c \in \mathcal{C}^*} E^{Q_x^{\varepsilon,c}}(f(x_\varepsilon(t)) - \xi_\varepsilon(t)) - \sup_{y \in \mathbb{R}^d}\left[f(y) - t\mathcal{I}\left(\frac{y - x}{t}\right)\right] \end{aligned}$$

$$\leqslant \sup_{c\in\mathcal{C}^*} E^{Q_x^{\varepsilon,c}}\left[\left[t\mathcal{I}\left(\frac{x_\varepsilon(t)-x}{t}\right)-\xi_\varepsilon(t)\right]\mathbf{1}_{\|x_\varepsilon(t)\|\leqslant M_\delta}\right]+\delta$$
$$=\sup_{c\in\mathcal{C}^*} E^{Q_x^{\varepsilon,c}}\left[\left(t\sup_{\theta\in\mathbb{R}^d}\left(\left\langle\theta,\frac{x_\varepsilon(t)-x}{t}\right\rangle-\overline{H}(\theta)\right)-\xi_\varepsilon(t)\right)\mathbf{1}_{\|x_\varepsilon(t)\|\leqslant M_\delta}\right]+\delta.$$

Note that according to (2.16),

$$\overline{H}(\theta)\geqslant c_1\|\theta\|^\alpha-c_2.$$

The above estimate together with estimate (7.1) ensures that large values of θ will not matter. In view of the continuity of $\overline{H}$, it suffices to show that for each fixed $\theta\in\mathbb{R}^d$, $\ell, T, \eta>0$,

$$\lim_{\varepsilon\to 0}\sup_{c\in\mathcal{C}^*}\sup_{\|x\|\leqslant\ell}\sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\left\{\langle\theta,x_\varepsilon(t)-x\rangle-t\overline{H}(\theta)-\xi_\varepsilon(t)\geqslant\eta\right\}=0. \tag{7.2}$$

Here we are using the facts that $\xi_\varepsilon(t)$ is bounded below uniformly in ε and that the family $\{\|x_\varepsilon\|\}$ is uniformly integrable.

PROOF OF THEOREM 2.2 ASSUMING (H3): According to Lemma 6.3, for any $\theta\in\mathbb{R}^d$ and $\delta>0$, there is a uniformly bounded vector field $\hat{v}_\delta$ that is a potential gradient. Let $V_\delta(x,\omega)$ be its integral as in Lemma 6.2. Then there exists a set N with $\mathbb{P}(N)=1$ such that for $\omega\in N$, $V_{\delta,\varepsilon}(x,\omega)=\varepsilon V_\delta(\varepsilon^{-1}x,\omega)$ satisfies

$$\lim_{\varepsilon\to 0}\sup_{\|x\|\leqslant\ell}|V_{\delta,\varepsilon}(x,\omega)|=0, \tag{7.3}$$

$$\sup_{0<\varepsilon\leqslant 1}|V_{\delta,\varepsilon}(x,\omega)|\leqslant M_\delta\|x\|, \tag{7.4}$$

for some constant M_δ. Applying Itô's formula to $V_{\delta,\varepsilon}$, a.e. with respect to $Q_x^{\varepsilon,c}$, we have

$$\begin{aligned}
&V_{\delta,\varepsilon}(x(t),\omega)-V_{\delta,\varepsilon}(x(0),\omega)\\
&\quad=\int_0^t\frac{1}{2}(\nabla\cdot\hat{v}_\delta)(\varepsilon^{-1}x(s),\omega)ds+\int_0^t\langle c(\varepsilon^{-1}s,\varepsilon^{-1}x(s)),\hat{v}_\delta(\varepsilon^{-1}x(s),\omega)\rangle ds\\
&\qquad+\sqrt{\varepsilon}\int_0^t\langle\hat{v}_\delta(\varepsilon^{-1}x(s),\omega),d\beta(s)\rangle.
\end{aligned}$$

Here and below, we simply write $x(t)$ for $x_\varepsilon(t)$. From (3.5) we also have

$$\langle\theta,x(t)-x\rangle=\int_0^t\langle\theta,c(\varepsilon^{-1}s,\varepsilon^{-1}x(s))\rangle ds+\sqrt{\varepsilon}\langle\theta,\beta(t)\rangle.$$

Adding the two, noticing that $\langle c,p\rangle\leqslant H(p,\omega)+L(c,\omega)$, and using the inequality of Lemma 6.3, we obtain

$$\begin{aligned}
&V_{\delta,\varepsilon}(x(t),\omega)-V_{\delta,\varepsilon}(x(0),\omega)+\langle\theta,x(t)-x\rangle\\
&\qquad\leqslant[\overline{H}(\theta)+\hat{\nu}(\delta)]\,t+\xi_\varepsilon(t)+\sqrt{\varepsilon}\int_0^t\langle\theta+\hat{v}_\delta(\varepsilon^{-1}x(s),\omega),d\beta(s)\rangle.
\end{aligned}$$

Since we consider only $c \in \mathcal{C}^*$, the family $\{\|x_\varepsilon(t)\|\}$ is uniformly integrable. This together with (7.3) and (7.4) imply

$$\lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{\|x\|\leqslant \ell} \sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\big[|V_{\delta,\varepsilon}(x(t)) - V_{\delta,\varepsilon}(x)| \geqslant \eta\big] = 0.$$

Since $\hat{v}_\delta \in L^\infty(\Omega, \mathbb{P})$,

$$\lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{\|x\|\leqslant \ell} \sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\left[\sqrt{\varepsilon}\left|\int_0^t \langle \theta + \hat{v}_\delta(\varepsilon^{-1}x(s), \omega), d\beta(s)\rangle\right| \geqslant \eta\right] = 0. \tag{7.5}$$

From $\lim_{\delta\to 0} \hat{\nu}(\delta) = 0$ and the last three inequalities, we obtain

$$\lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{\|x\|\leqslant \ell} \sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\big[\langle \theta, x(t) - x\rangle - \xi_\varepsilon(t) - t\overline{H}(\theta) \geqslant \eta\big] = 0,$$

thereby completing the proof of Theorem 2.2 assuming (H3). □

The main problem in the absence of assumption (H3) is that we cannot regularize v to get a bounded v_δ that still satisfies the inequality of Lemma 6.3. In the absence of regularization, one cannot get a bound $\|\nabla v\|_\infty$, which is needed to control the stochastic integral term (7.5) as well as to obtain the uniformity in the ergodic theorem. This is handled differently in the two cases depending on whether (H4) or (H5) is assumed.

PROOF OF THEOREM 2.2 ASSUMING (H4): In this case since $v \in L^\alpha(\mathbb{P})$ with $\alpha > d$, one uses Sobolev's imbedding to control the $\|V\|_\infty$ locally and then martingale theory to control, also locally, the quadratic variation

$$E^{Q_x^{\varepsilon,c}}\left[\int_0^\tau \|v(\varepsilon^{-1}x(s))\|^2\, ds\right]$$

for a suitable stopping time τ.

Recall $v(x, \omega) = v(\tau_x\omega)$. Since $v(\,\cdot\,, \omega) \in L^\alpha_{\mathrm{loc}}(\mathbb{R}^d)$ for almost all ω, Sobolev's imbedding theorem gives us an almost surely continuous choice of $V(x, \omega)$ that can be normalized so that $V(0, \omega) = 0$ for $\mathbb{P}$-a.e. ω.

Step 1. Let $\tau_{\varepsilon,a} = \inf\{t > 0 : \xi_\varepsilon(t) \geqslant a\}$, $\hat{\tau}_{\varepsilon,b} = \inf\{t > 0 : \|x_\varepsilon(t) - x\| \geqslant b\}$, and $\sigma = \sigma(\varepsilon, a, b) = \tau_{\varepsilon,a} \wedge \hat{\tau}_{\varepsilon,b}$. Then an argument similar to the proof of Lemma 4.1 shows that for sufficiently large a and b

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{\|x\|\leqslant \ell} \sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\big\{\langle \theta, x(t) - x\rangle - \overline{H}(\theta)t - \xi_\varepsilon(t) \geqslant \eta\big\}\\
&\quad = \lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{\|x\|\leqslant \ell} \sup_{0\leqslant t\leqslant T} Q_x^{\varepsilon,c}\big\{\langle \theta, x(t) - x\rangle - \overline{H}(\theta)t - \xi_\varepsilon(t) \geqslant \eta;\ \sigma \geqslant t\big\}
\end{aligned}$$

Let $V_{\delta,\varepsilon}(x,\omega)=\varepsilon V_\delta(x/\varepsilon,\omega)$. Recalling definitions (3.5) and (3.6) and applying the rescaled version of Lemma 6.4, we get for each $c\in\mathcal{C}$ and $\varepsilon>0$

$$\begin{aligned}&\langle\theta,x(t)-x\rangle-t\overline{H}(\theta)-\xi_\varepsilon(t)\\&\quad\leqslant V_{\delta,\varepsilon}(x,\omega)-V_{\delta,\varepsilon}(x(t),\omega)\\&\qquad+\sqrt{\varepsilon}\int_0^t\langle\theta+v_\delta(\tau_{\varepsilon^{-1}x(s)}\omega),d\beta(s)\rangle+r_c(\delta)t\end{aligned}\tag{7.6}$$

where $r_c(\delta)\to 0$ as $\delta\to 0$ for each $c\in\mathcal{C}$. An application of Itô's formula gives

$$\begin{aligned}&\sqrt{\varepsilon}\int_0^t\langle v_\delta(\tau_{x(s)}\omega),d\beta(s)\rangle\\&=V_{\delta,\varepsilon}(x(t),\omega)-V_{\delta,\varepsilon}(x,\omega)\\&\quad-\int_0^t\frac{\varepsilon}{2}\nabla\cdot v_\delta(\tau_{\varepsilon^{-1}x(s)}\omega)+\langle c(s,x(s)),v_\delta(\tau_{\varepsilon^{-1}x(s)}\omega)\rangle\,ds.\end{aligned}\tag{7.7}$$

Substituting this into (7.6) we obtain, almost surely with respect to $Q_x^{\varepsilon,c}$,

$$\begin{aligned}&\langle\theta,x(t)-x\rangle-t\overline{H}(\theta)-\xi_\varepsilon(t)\\&\quad\leqslant\sqrt{\varepsilon}\langle\theta,\beta(t)\rangle\\&\qquad-\int_0^t\frac{\varepsilon}{2}\nabla\cdot v_\delta(\tau_{\varepsilon^{-1}x(s)}\omega)+\langle c(s,x(s)),v_\delta(\tau_{\varepsilon^{-1}x(s)}\omega)\rangle ds+r_c(\delta)t.\end{aligned}$$

Setting

$$F_{\delta,\varepsilon}(s,x,\omega)=\frac{\varepsilon}{2}\nabla\cdot v_\delta(\tau_{\varepsilon^{-1}x}\omega)+\langle c(s,x),v_\delta(\tau_{\varepsilon^{-1}x}\omega)\rangle,$$

we observe that it is enough to show that for almost all ω

$$\lim_{\varepsilon\to0}\sup_{c\in\mathcal{C}^*}\sup_{\|x\|\leqslant\ell}\sup_{0\leqslant t\leqslant T}\limsup_{\delta\to0}E^{Q_x^{\varepsilon,c}}\left(\int_0^{t\wedge\sigma}F_{\delta,\varepsilon}(s,x(s),\omega)ds\right)^2=0.\tag{7.8}$$

Step 2. Below we shall write E_x instead of $E^{Q_x^{\varepsilon,c}}$ and also drop ω from the notation. Rewriting the square of the integral as twice the time-ordered double integral, we get

$$\begin{aligned}&E_x\left(\int_0^{t\wedge\sigma}F_{\delta,\varepsilon}(s,x(s))ds\right)^2\\&\qquad=2E_x\int_0^{t\wedge\sigma}F_{\delta,\varepsilon}(s,x(s))\left(\int_s^{t\wedge\sigma}F_{\delta,\varepsilon}(u,x(u))du\right)ds.\end{aligned}$$

After conditioning on the σ-algebra up to time s and using (7.7), we see that the last integral is equal to

$$2E_x\int_0^{t\wedge\sigma}F_{\delta,\varepsilon}(s,x(s))E_{x(s)}\big(V_{\delta,\varepsilon}(x(t\wedge\sigma))-V_{\delta,\varepsilon}(x(s))\big)ds.$$

Thus, for all starting points x with $\|x\| \leqslant \ell$

$$\begin{aligned}
&E_x\left[\int_0^{t\wedge\sigma} F_{\delta,\varepsilon}(s, x(s))ds\right]^2 \\
&\quad\leqslant 4 \sup_{\|x\|\leqslant b+\ell} |V_{\delta,\varepsilon}(x)| E_x \int_0^{t\wedge\sigma} |F_{\delta,\varepsilon}(s, x(s))|ds \\
&\quad= 4 \sup_{\|x\|\leqslant b+\ell} |V_{\delta,\varepsilon}(x)| \left[2E_x \int_0^{t\wedge\sigma} [F_{\delta,\varepsilon}(s, x(s))]_+ \, ds \right.\\
&\qquad\qquad\qquad\qquad\qquad \left. - E_x \int_0^{t\wedge\sigma} F_{\delta,\varepsilon}(s, x(s))ds\right] \\
&\quad\leqslant 8 \sup_{\|x\|\leqslant b+\ell} |V_{\delta,\varepsilon}(x)| \left[E_x \int_0^{t\wedge\sigma} [F_{\delta,\varepsilon}(s, x(s))]_+ \, ds + \sup_{\|x\|\leqslant b+\ell} |V_{\delta,\varepsilon}(x)|\right].
\end{aligned}$$

Step 3. Only now we use the assumption $\alpha > d$. Sobolev's imbedding theorem and Lemma 6.1 imply that for almost all ω, the function V satisfies

$$|V(x, \omega) - V(y, \omega)| \leqslant C(\|x - y\|),$$

where $C(r) \to 0$ as $r \to 0$ and $C(r) \leqslant Cr$ for $r \geqslant 1$. By the uniform continuity and the ergodic theorem there is a set $N \subset \Omega$ with $\mathbb{P}(N) = 1$ such that for all $\omega \in N$

$$\lim_{\varepsilon\to 0} \lim_{\delta\to 0} V_{\delta,\varepsilon}(x, \omega) = 0$$

uniformly on bounded subsets of $\mathbb{R}^d$. Thus, a bound on

$$E_x \int_0^{t\wedge\sigma} [F_{\delta,\varepsilon}(s, x(s))]_+ \, ds$$

will complete the proof. From the definition of $F_{\delta,\varepsilon}$, the rescaled versions of (6.7) and (6.3), and the weak regularity of L, we obtain

$$\begin{aligned}
E_x \int_0^{t\wedge\sigma} [F_{\delta,\varepsilon}(s, x)]_+ \, ds \leqslant E_x \int_0^{t\wedge\sigma} &\left[\overline{H}(\theta) + L(\varepsilon^{-1}s, \varepsilon^{-1}x(s)), \tau_{\varepsilon^{-1}x(s)}\omega)\right. \\
&\left. - \langle\theta, c(\varepsilon^{-1}s, \varepsilon^{-1}x(s))\rangle + r_c(\delta)\right]_+ ds
\end{aligned}$$

as $\delta \to 0$. Lemma 4.1 implies the desired bound (7.8) and finishes the proof. □

Finally, we turn to the proof of Theorem 2.2 under assumptions (H1), (H2), and (H5). First we prove almost sure weak convergence of $u_\varepsilon(t, x, \omega)$ to $u(t, x)$. Looking at (7.6) we observe that the main problem is getting a control on the martingale term. The idea of the proof is to perturb the Hamiltonian H by replacing it with $H^\gamma(p, \omega) = H(p, \omega) + \gamma\|p\|^2$ with small $\gamma > 0$ and then using the "leftover" quadratic term to control the martingale by means of the following lemma (see, for example, section 2.3 of [14]).

LEMMA 7.1 *Let z be a nonanticipating Brownian functional such that*

$$P\left[\int_0^t z^2\,dt < \infty, t \geqslant 0\right] = 1.$$

Then

$$P\left[\sup_{t\geqslant 0}\left(\int_0^t z\,d\beta - \frac{a}{2}\int_0^t z^2\,ds\right) > b\right] \leqslant e^{-ab}.$$

After passing to the limit as $\varepsilon \to 0$ we would let $\gamma \to 0$. Assumption (H5) will assure that the new effective Hamiltonian $\overline{H}^{\gamma}$ decreases to $\overline{H}^{0} = \overline{H}$ as $\gamma \downarrow 0$.

LEMMA 7.2 *For any $\theta \in \mathbb{R}^d$*

$$\lim_{\gamma\to 0} \overline{H}^{\gamma}(\theta) = \overline{H}(\theta).$$

PROOF: Obviously, $\overline{H}^{\gamma_1} \geqslant \overline{H}^{\gamma_2}$ for $\gamma_1 \geqslant \gamma_2 \geqslant 0$ and $\lim_{\gamma\to 0} \overline{H}^{\gamma}(\theta) \geqslant \overline{H}(\theta)$. We need to prove that $\lim_{\gamma\to 0} \overline{H}^{\gamma}(\theta) \leqslant \overline{H}(\theta)$.

Fix $\theta \in \mathbb{R}^d$ and apply Theorem 5.2 to get the gradient v of a supersolution corresponding to $\theta/(1-\delta)$. Then the function $\psi = (1-\delta)v$ satisfies

$$\frac{1}{2}(\nabla\cdot\psi)(\omega) + (1-\delta)H\left(\frac{\theta+\psi(\omega)}{1-\delta}, \omega\right) \leqslant (1-\delta)\overline{H}\left(\frac{\theta}{1-\delta}\right).$$

From (H5) we get

$$(1-\delta)H\left(\frac{\theta+p}{1-\delta}, \omega\right) \geqslant H(\theta+p,\omega) + \gamma(\delta)\|\theta+p\|^2 - C(\delta).$$

Then for all $0 \leqslant \gamma \leqslant \gamma(\delta)$

$$\frac{1}{2}(\nabla\cdot\psi)(\omega) + H^{\gamma}(\theta+\psi(\omega),\omega) \leqslant (1-\delta)\overline{H}\left(\frac{\theta}{1-\delta}\right) + C(\delta). \tag{7.9}$$

Using Corollary 6.9 with

$$\lambda = (1-\delta)\overline{H}\left(\frac{\theta}{1-\delta}\right) + C(\delta)$$

and $\overline{H}(\theta)$ replaced by $\overline{H}^{\gamma}(\theta)$, we obtain the estimate

$$\overline{H}^{\gamma}(\theta) \leqslant (1-\delta)\overline{H}\left(\frac{\theta}{1-\delta}\right) + C(\delta).$$

Finally, we let $\gamma \to 0$ and then $\delta \to 0$. □

Let v_δ and V_δ be convolutions with a mollification kernel $\rho_\delta(x) = \delta^{-d}\rho(x/\delta)$ as in Lemma 6.2.

PROOF OF OF THEOREM 2.2 UNDER (H5): The proof in this case involves several steps.

Step 1. Recall that Q_x^c denotes the law of a diffusion with drift c that starts from a point x. Let us write Q_φ^c for the law of the same diffusion when the initial position is distributed according to a probability measure with a density φ. We start with $H^\gamma(p,\omega) = H(p,\omega)+\gamma\|p\|^2$ and construct the corresponding v^γ using Theorem 5.2. We have that $\mathbb{E}\, v^\gamma = 0$, $\nabla \times v^\gamma = 0$, and

$$\frac{1}{2}\nabla \cdot v^\gamma + H(\theta + v^\gamma, \omega) + \gamma\|v^\gamma\|^2 \leqslant \overline{H}^\gamma(\theta).$$

Then we define V^γ as in (6.1) so that $V^\gamma \in L^\alpha_{\mathrm{loc}}(\mathbb{R}^d)$, $\nabla V^\gamma = v^\gamma$, and $V^\gamma(0,\omega) = 0$ for $\mathbb{P}$-a.e. ω. They are then smoothed to obtain V_δ^γ satisfying $\nabla V_\delta^\gamma = v_\delta^\gamma$ and $V_\delta^\gamma(x) = \int V^\gamma(x+\delta y,\omega)\rho(y)dy$. Let $c \in \mathcal{C}$. Then by Itô's formula,for $\mathbb{P}$-a.e. ω and Q_φ^c-a.e. path $x(\cdot)$

$$\begin{aligned}
&V_\delta^\gamma((x(t)) - V_\delta^\gamma(x(0)) + \langle\theta, x(t) - x(0)\rangle \\
&\quad = \int_0^t \left(\frac{1}{2}\nabla \cdot v_\delta^\gamma + c\cdot(\theta + v_\delta^\gamma)\right)(s, x(s))ds + \int_0^t \langle (v_\delta^\gamma((x(s)) + \theta), d\beta(s)\rangle.
\end{aligned}$$

We add $\gamma\int_0^t \|\theta + v_\delta^\gamma\|^2\, ds$ to both sides and use the inequality

$$\langle c, p\rangle \leqslant L(c,\omega) + H^\gamma(p,\omega) - \gamma\|p\|^2$$

to obtain

$$\begin{aligned}
&\langle\theta, x(t) - x(0)\rangle + \gamma\int_0^t \|\theta + v_\delta^\gamma(x(s))\|^2\, ds \\
&\quad - \int_0^t L(c(s,x(s)), \tau_{x(s)}\omega)ds - t\overline{H}^\gamma(\theta) \\
&\qquad \leqslant -V_\delta^\gamma((x(t)) + V_\delta^\gamma(x(0)) + \int_0^t \langle v_\delta^\gamma(x(s)) + \theta, d\beta(s)\rangle + r_c(\delta)t,
\end{aligned}$$

Q^c_φ-a.e. as in Lemma 6.4. If the initial distribution of $x(0)$ were given by a nice density $\varphi(x)dx$, then since t and $c \in \mathcal{C}$ are fixed, we could let $\delta \to 0$ to arrive at

$$\begin{aligned}
&\langle \theta, x(t) - x(0)\rangle - \int_0^t L(c(s, x(s)), \tau_{x(s)}\omega)ds - t\overline{H}^\gamma(\theta) \\
&\leqslant -V^\gamma((x(t)) + V^\gamma(x(0)) + \int_0^t \langle v^\gamma(x(s)) + \theta, d\beta(s)\rangle \\
&\quad - \gamma \int_0^t \|\theta + v^\gamma(x(s))\|^2\, ds
\end{aligned}$$

a.e. Q^c_φ.

We can now use Lemma 6.5 to replace $V^\gamma(x(t))$ by $V^\gamma_\delta(x(t))$:

$$\begin{aligned}
&\langle \theta, x(t) - x(0)\rangle - \int_0^t L(c(s, x(s)), \tau_{x(s)}\omega)ds - t\overline{H}^\gamma(\theta) \\
&\leqslant -V^\gamma_\delta((x(t)) + V^\gamma(x(0)) + \int_0^t \langle v^\gamma(x(s)) + \theta, d\beta(s)\rangle \\
&\quad - \gamma \int_0^t \|\theta + v^\gamma(x(s))\|^2\, ds + C\delta^2
\end{aligned}$$

a.e. Q^c_φ. Now we rescale and let $\varepsilon \to 0$, allowing arbitrary choices for $c = c_\varepsilon \in \mathcal{C}^*$ and suitable choices of $\varphi = \varphi_\varepsilon$. In fact, the choices of c_ε will be limited to those that satisfy

$$E^{Q^{c_\varepsilon}_\varphi}\left[\gamma \int_0^t \|\theta + v^\gamma(x(s))\|^2\, ds\right] \leqslant Ct$$

for some C independent of ε. By the ergodic theorem, $\varepsilon V^\gamma_\delta(x(\varepsilon^{-1}t)) \to 0$. By Lemma 7.1 the stochastic integral term goes to 0. If the choice of $\varphi_\varepsilon(y)$ is of the form $\varphi(x_\varepsilon - y)$, with fixed φ and $\varepsilon x_\varepsilon \to x$, then the ergodic theorem again can be used to show that

$$\lim_{\varepsilon\to 0} E^{Q^{c_\varepsilon}_{\varphi_\varepsilon}}[\varepsilon|V^\gamma(x(0))|] = 0.$$

Now one can see that with probability 1

$$\lim_{\varepsilon\to 0} \sup_{0\leqslant t\leqslant T} \sup_{|x|\leqslant \ell} \left| \int u_\varepsilon(t, x + \varepsilon y, \omega)\varphi(y)dy - u(t, x)\right| = 0. \tag{7.10}$$

Step 2. By (H5) the function L satisfies

$$L(q, \omega) \leqslant c_3\|q\|^{\alpha'} + c_4$$

with $\alpha' < 2$. Then it is possible to pick $c(\,\cdot\,,\cdot\,) : [0, t] \times \mathbb{R}^d \to \mathbb{R}^d$ such that the process Q^c_x is the Brownian bridge from x to y with $Q^c_x[x(t) = y] = 1$. For a Brownian bridge the choice of c is

$$c(s, x) = \frac{y - x}{t - s}.$$

Since

$$x(s) = \left(1 - \frac{s}{t}\right)x + \frac{s}{t}y + \beta(s) - \frac{s}{t}\beta(t), \quad 0 \leqslant s \leqslant t,$$

we have that

$$c(s, x(s)) = \frac{y - x(s)}{t - s} = \frac{y - x}{t} + \frac{\beta(t) - \beta(s)}{t - s} - \frac{\beta(t)}{t}$$

and

$$\|c(s, x(s))\|^{\alpha'} \leqslant A\left\|\frac{y - x}{t}\right\|^{\alpha'} + A\left\|\frac{\beta(t) - \beta(s)}{t - s}\right\|^{\alpha'} + A\left\|\frac{\beta(t)}{t}\right\|^{\alpha'}$$

with $A = 3^{\alpha'-1}$. Computing the expected value, we get

$$E^{Q_x^c}\|c(s, x(s))\|^{\alpha'} \leqslant A\left\|\frac{y - x}{t}\right\|^{\alpha'} + \frac{B}{(t - s)^{\alpha'/2}} + \frac{B}{t^{\alpha'/2}}.$$

For $\alpha' < 2$ the integration from 0 to t gives

$$E^{Q_x^c}\left[\int_0^t \|c(s, x)\|^{\alpha'} ds\right] \leqslant At\left\|\frac{|x - y|}{t}\right\|^{\alpha'} + Ct^{1-\alpha'/2} < \infty,$$

providing the following lower bound for $v_\varepsilon(t, x, \omega)$, the solution of (3.1):

$$v_\varepsilon(t, x, \omega) \geqslant \sup_y \left[\frac{1}{\varepsilon}f(\varepsilon y) - At\left\|\frac{|x - y|}{t}\right\|^{\alpha'} - Ct^{1-\alpha'/2} - c_4 t\right].$$

Recalling that $u_\varepsilon(t, x, \omega) = \varepsilon v_\varepsilon(t/\varepsilon, x/\varepsilon, \omega)$, we obtain

$$\begin{aligned} u_\varepsilon(t, x, \omega) &\geqslant \sup_y \left[f(\varepsilon y) - At\left\|\frac{|x - \varepsilon y|}{t}\right\|^{\alpha'} - \varepsilon C\left(\frac{t}{\varepsilon}\right)^{1-\alpha'/2} - c_4 t\right] \\ &= \sup_y \left[f(y) - At\left\|\frac{|x - y|}{t}\right\|^{\alpha'} - C\varepsilon^{\alpha'/2}t^{1-\alpha'/2} - c_4 t\right]. \end{aligned}$$

Combining the above estimate with the semigroup property, we arrive at

$$u_\varepsilon(t+h, x, \omega) \geqslant \sup_y \left[u_\varepsilon(t, y, \omega) - Ah^{1-\alpha'}\|x - y\|^{\alpha'} - C\varepsilon^{\alpha'/2}h^{1-\alpha'/2} - c_4 h\right]. \tag{7.11}$$

Step 3. It follows from (7.10) that uniformly for $(t, x) \in [0, T] \times \{x : \|x\| \leqslant \ell\}$ and $0 \leqslant h \leqslant 1$

$$\limsup_{\varepsilon \to 0} \inf_{y:\|y - x\| \leqslant \varepsilon} u_\varepsilon(t + h, y, \omega) \leqslant u(t + h, x).$$

Then from (7.11) we obtain

$$\inf_{y:\|y - x\| \leqslant \varepsilon} u_\varepsilon(t + h, y, \omega) \geqslant u_\varepsilon(t, x, \omega) - Ah^{1-\alpha'}\varepsilon^{\alpha'} - C\varepsilon^{\alpha'/2} - c_4 h.$$

The above two inequalities imply the estimate

$$\limsup_{\varepsilon \to 0} u_\varepsilon(t, x, \omega) \leqslant u(t + h, x) + c_4 h$$

uniformly for $(t, x) \in [0, t) \times \{x : \|x\| \leqslant \ell\}$. Since $u(t, x)$ is given explicitly by (2.14), one can now let $h \to 0$ and obtain the upper bound. $\square$

Remark 7.3. If (H5) holds with $\alpha = 2$, then we only get (7.10). This, of course, implies local L^p convergence of u_ε to u.

Remark 7.4. In the absence of (H3), (H4), and (H5), one cannot control the stochastic integral, and one can only prove (7.10) for linear f. If $\alpha > 2$, this convergence can be improved to locally uniform convergence, as we have shown above.

Acknowledgment. Kosygina was partially supported by an NSF-AWM Mentoring Travel Grant and PSC-CUNY Awards #60079-33-34 and #66512-00-35. Rezakhanlou and Varadhan were partially supported, respectively, by NSF grants DMS-0307021 and DMS-0104343.

Bibliography

[1] Bensoussan, A.; Lions, J.-L.; Papanicolaou, G. *Asymptotic analysis for periodic structures.* Studies in Mathematics and Its Applications, 5. North-Holland, Amsterdam-New York, 1978.

[2] Braides, A.; Defranceschi, A. *Homogenization of multiple integrals.* Oxford Lecture Series in Mathematics and Its Applications, 12. The Clarendon Press, New York, 1998.

[3] Caffarelli, L. A.; Souganidis, P. E.; Wang, L. Stochastic homogenization for fully nonlinear, second order partial differential equations. *Comm. Pure Appl. Math.* **58** (2005), no. 3, 319–361.

[4] Capuzzo-Dolcetta, I.; Ishii, H. On the rate of convergence in homogenization of Hamilton-Jacobi equations. *Indiana Univ. Math. J.* **50** (2001), no. 3, 1113–1129.

[5] Cioranescu, D.; Donato, P. *An introduction to homogenization*. Oxford Lecture Series in Mathematics and Its Applications, 17. The Clarendon Press, New York, 1999.

[6] Concordel, M. C. Periodic homogenization of Hamilton-Jacobi equations: additive eigenvalues and variational formula. *Indiana Univ. Math. J.* **45** (1996), no. 4, 1095–1117.

[7] Dal Maso, G. *An introduction to Γ-convergence*. Progress in Nonlinear Differential Equations and Their Applications, 8. Birkhäuser-Boston, Boston, 1993.

[8] Evans, L. Periodic homogenisation of certain fully nonlinear partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A* **120** (1992), no. 3-4, 245–265.

[9] Fleming, W. H.; Soner, H. M. *Controlled Markov processes and viscosity solutions.* Applications of Mathematics (New York), 25. Springer, New York, 1993.

[10] Jikov, V. V.; Kozlov, S. M.; Oleĭnik, O. A. *Homogenization of differential operators and integral functionals*. Translated from the Russian by G. A. Yosifian. Springer, Berlin, 1994.

[11] Lions, P. -L.; Papanicolaou, G.; Varadhan, S. R. S. Homogenization of Hamilton-Jacobi equations. Unpublished manuscript, 1986.

[12] Lions, P. -L.; Souganidis, P. E. Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting. *Comm. Pure Appl. Math.* **56** (2003), no. 10, 1501–1524.

[13] Lions, P. -L.; Souganidis, P. E. Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media. Preprint, 2004.

[14] McKean, H. P. Jr. *Stochastic integrals*. Probability and Mathematical Statistics, 5. Academic Press, New York, 1969.

[15] Pankov, A. *G-convergence and homogenization of nonlinear partial differential operators*. Mathematics and Its Applications, 422. Kluwer, Dordrecht, 1997.

[16] Papanicolaou, G. C.; Varadhan, S. R. S. *Diffusions with random coefficients.* North-Holland, Amsterdam, 1982.

[17] Rezakhanlou, F.; Tarver, J. E. Homogenization for stochastic Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.* **151** (2000), no. 4, 277–309.

[18] Sion, M. On general minimax theorems. *Pacific J. Math.* **8** (1958), 171–176.

[19] Souganidis, P. Stochastic homogenization of Hamilton-Jacobi equations and some applications. *Asymptot. Anal.* **20** (1999), no. 1, 1–11.

ELENA KOSYGINA
Department of Mathematics
Baruch College
One Bernard Baruch Way
New York, NY 10010
Email: elena_kosygina@baruch.cuny.edu

FRAYDOUN REZAKHANLOU
Department of Mathematics
University of California, Berkeley
815 Evans Hall
Berkeley, CA 94720
Email: rezakhan@math.berkeley.edu

S. R. S. VARADHAN
Courant Institute
251 Mercer St.
New York, NY 10012
Email: varadhan@cims.nyu.edu

Received January 2005.