Homogenization of Hamilton-Jacobi-Bellman Equations with Respect to Time-Space Shifts in a Stationary Ergodic Medium

ELENA KOSYGINA
Baruch College
CUNY Graduate Center

AND

S. R. S. VARADHAN
Courant Institute

Abstract

We consider a family $\{u_\varepsilon(t,x,\omega)\}$, $\varepsilon > 0$, of solutions to the equation $\partial u_\varepsilon/\partial t + \varepsilon\Delta u_\varepsilon/2 + H(t/\varepsilon, x/\varepsilon, \nabla u_\varepsilon, \omega) = 0$ with the terminal data $u_\varepsilon(T,x,\omega) = U(x)$. Assuming that the dependence of the Hamiltonian $H(t,x,p,\omega)$ on time and space is realized through shifts in a stationary ergodic random medium, and that H is convex in p and satisfies certain growth and regularity conditions, we show the almost sure locally uniform convergence, in time and space, of $u_\varepsilon(t,x,\omega)$ as $\varepsilon \to 0$ to the solution $u(t,x)$ of a deterministic averaged equation $\partial u/\partial t + \overline{H}(\nabla u) = 0$, $u(T,x) = U(x)$. The "effective" Hamiltonian $\overline{H}$ is given by a variational formula.

1 Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\omega \in \Omega$. We are interested in the behavior, as $\varepsilon \to 0$, of a family $\{u_\varepsilon(t,x,\omega)\}$, $\varepsilon > 0$, of solutions to the following terminal value problem:

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

$$u_\varepsilon(T,x,\omega) = U(x). \tag{1.2}$$

The Hamiltonian $H(t,x,p,\omega)$ is assumed to be convex in p. The dependence of H on (t,x) is realized through the shifts in a stationary ergodic random medium (see the next section for a detailed description). The terminal condition U is a uniformly continuous nonrandom function. Under some additional assumptions on H, we obtain a homogenization result; i.e., we show that, with probability 1, $u_\varepsilon(t,x,\omega)$ converges locally uniformly in t and x to a nonrandom limit $u(t,x)$.

Function $u(t,x)$ is the unique solution of the Hamilton-Jacobi equation

$$\frac{\partial u}{\partial t} + \overline{H}(\nabla u) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^d, \tag{1.3}$$

with the same terminal value $U(x)$. We also provide a variational representation for the effective Hamiltonian $\overline{H}$.

Homogenization problems for this type of equation with or without a "viscous" term were extensively studied in the periodic, quasi, and almost periodic settings. We mention only several relatively recent papers where one can find further references [1, 2, 6, 8, 10, 17]. For a broader view of homogenization methods, see monographs [3, 4, 5, 9, 13, 19].

One of the standard approaches is based on the construction of so-called correctors. Roughly speaking, the corrector is the first nontrivial error term in a formal asymptotic expansion of u_ε around u. It is obtained as a solution of an auxiliary "cell" problem. The word *cell* refers to the fact that in the periodic case the space variable changes over a torus. The existence of correctors and the uniform convergence of u_ε to a limit were established for a wide range of first- and second-order partial differential equations. This method is very robust and does not require the Hamiltonian to be convex. The most essential assumption on the Hamiltonian is the coercivity in p:

$$H(t,x,p,\omega) \to \infty \quad \text{as } |p| \to \infty,$$

uniformly in t, x, and ω.

The extension of homogenization results to the stationary ergodic setting presented a number of difficulties due to the lack of compactness. It was shown that correctors (or even approximate correctors) need not exist in general [16]. In the case when H is convex in p and independent of t, the homogenization can be proved by using variational methods in combination with some version of the ergodic theorem (see [14, 18]). A further step was taken in [7], which proves homogenization for fully nonlinear uniformly elliptic equations whose solutions do not have a representation formula.

It turns out that the time-space averaging (versus just the space averaging) requires an additional control on u_ε (see, for example, [12] for periodic Hamilton-Jacobi equations and [15] for linear equations in stationary ergodic random media).

In this paper we present a method that allows us to obtain time-space homogenization results for Hamilton-Jacobi-Bellman equations in a stationary ergodic setting. It is essential for our approach that H be convex in p.

In the case when

$$H(t,x,p,\omega) = \frac{1}{2}\sum_{i=1}^{d} p_i^2 - \sum_{i=1}^{d} b_i(t,x,\omega)p_i - W(t,x,\omega),$$

the homogenization problem for (1.1) is closely related to the quenched large-deviations principle for a Brownian motion with a random drift $b = (b_1, b_2,$

$\dots, b_d)$ in a random potential W (see [14, 18, 21]). Both the drift and the potential are assumed to be stationary ergodic processes (see Remark 2.1).

The paper is organized as follows: In Section 2 we state the assumptions, formulate the main result, and discuss the idea of the proof. In Section 3 we prove the lower bound on u_ε. The construction of approximate supersolutions of (1.1) needed for the upper bound is carried out in Section 4. The proof of the upper bound on u_ε is given in Section 5. A generalization of Wiener's ergodic theorem used for the upper bound is derived in Section 6. The appendix contains proofs of two technical lemmas.

Notation. For $x, y \in \mathbb{R}^k$ we denote by $\langle x, y\rangle$ the standard scalar product of x and y and by $|y|$ the Euclidean norm of y. Borel σ-algebra on $\mathbb{R}^{d+1}$ is denoted by $\mathcal{B}$. If A is a set in $\mathbb{R}^k$, then $|A|$ denotes the Lebesgue measure of A. $B(x,r) \subset \mathbb{R}^k$ is a ball of radius r centered at x. If $f \in L^\alpha$, then $\|f\|_\alpha$ stands for its norm. $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$. $B(t)$, $t \geqslant 0$, is the path of the canonical Brownian motion.

2 Main Result and Preliminary Discussion

Let $\{\tau_{(t,x)} : (t,x) \in \mathbb{R}^{d+1}\}$ be a group of measure-preserving transformations acting ergodically on $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that the map $(t, x, \omega) \mapsto \tau_{(t,x)}\omega$ from $\mathbb{R}^{d+1} \times \Omega$ to Ω is $\mathcal{B} \times \mathcal{F}$ measurable.

Let $H(p,\omega) : \mathbb{R}^d \times \Omega \to \mathbb{R}$ be convex in p and satisfy the following assumptions:

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

$$c_1|p|^\alpha - c_2 \leqslant H(p,\omega) \leqslant c_3(|p|^\beta + 1)$$

for some positive constants c_1, c_2, c_3 and $1 < \alpha \leqslant \beta < \infty$.

(H2) $H(p, \tau_{(t,x)}\omega)$ is uniformly continuous in (t,x), uniformly in ω, and locally uniformly in p, i.e., for every $l > 0$

$$\lim_{\delta\to 0} \sup_{|(t,x)|\leqslant\delta} \sup_{|p|\leqslant l} \sup_{\omega\in\Omega} |H(p, \tau_{(t,x)}\omega) - H(p,\omega)| = 0.$$

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

$$H(p, \tau_{(t,x)}\omega) \geqslant (1+\nu(\delta))H((1+\nu(\delta))^{-1}p, \omega) - C\nu(\delta).$$

These assumptions can be equivalently stated in terms of the Lagrangian

$$L(q,\omega) = \sup_{p\in\mathbb{R}^d} (\langle p, q\rangle - H(p,\omega)). \tag{2.1}$$

(L1) For all $(p,\omega) \in \mathbb{R}^d \times \Omega$

$$c_4|q|^{\beta'} - c_5 \leqslant L(q,\omega) \leqslant c_6(|q|^{\alpha'} + 1)$$

for some positive of constants c_4, c_5, c_6 and

$$\alpha' = \frac{\alpha}{\alpha - 1}, \quad \beta' = \frac{\beta}{\beta - 1}.$$

(L2) $L(q, \tau_{(t,x)}\omega)$ is uniformly continuous in x, uniformly in ω, and locally uniformly in q.

(L3) There exists a positive function $\nu(\delta) \to 0$ as $\delta \to 0$ and a constant $C > 0$ such that for $|(t, x)| \leqslant \delta$ and $\omega \in \Omega$

$$L(q, \tau_{(t,x)}\omega) \leqslant (1 + \nu(\delta))L(q, \omega) + C\nu(\delta).$$

We assume that the terminal data $U(x)$ is uniformly continuous. Note that such U can be uniformly approximated by functions that are Lipschitz-continuous, and by the maximum principle the u_ε are uniformly approximated as well. So there is no loss of generality in assuming that there is a constant K such that for all $x, y, \in \mathbb{R}^d$

$$|U(x) - U(y)| \leqslant K|x - y|. \tag{2.2}$$

This makes the proof somewhat simpler.

Remark 2.1. Functions $H(t, x, p, \omega)$, $b(t, x, \omega)$, and $W(t, x, \omega)$, from the introduction, should be understood as $H(p, \tau_{(t,x)}\omega)$, $b(\tau_{(t,x)}\omega)$, and $W(\tau_{(t,x)}\omega)$, respectively.

2.1 The Effective Hamiltonian $\overline{H}$

The translation group $\{\tau_{(t,x)} : (t, x) \in \mathbb{R}^{d+1}\}$ acting on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ will have infinitesimal generators D_t, ∇_i, $i = 1, 2, \dots, d$, in the coordinate directions (see [13, sec. 7.1 and pp. 231–232]). The space gradient $\nabla = (\nabla_1, \nabla_2, \dots, \nabla_d)$, the divergence $\operatorname{div}(U_1, U_2, \dots, U_d) = \nabla_1 U_1 + \nabla_2 U_2 + \cdots \nabla_d U_d$, and the Laplace operator $\Delta = \sum_{i=1}^d \nabla_i^2$ are defined in the usual way.

Let us denote by $\mathbf{B}$ the space of measurable, essentially bounded maps from Ω to $\mathbb{R}^d$ and by $\mathbf{D}$ the space of bounded probability densities $\Phi : \Omega \to \mathbb{R}$ relative to $\mathbb{P}$ that are bounded away from 0 and have essentially bounded time-space gradients. Define

$$\mathcal{E} = \{(b, \Phi) \in \mathbf{B} \times \mathbf{D} : D_t\Phi + \operatorname{div}(b\Phi) = \tfrac{1}{2}\Delta\Phi\}. \tag{2.3}$$

We shall always assume that the equation in (2.3) is satisfied in the weak sense: with probability 1 for every $G \in C_0^\infty(\mathbb{R}^{d+1})$

$$\iint \Phi(\tau_{(t,x)}\omega)\partial_t G(t, x) + \left\langle (b\Phi - \tfrac{1}{2}\nabla\Phi)(\tau_{(t,x)}\omega), \nabla G(t, x)\right\rangle dt\, dx = 0. \tag{2.4}$$

For $(b, \Phi) \in \mathcal{E}$ set

$$m(b, \Phi) = \mathbb{E}[b(\omega)\Phi(\omega)], \tag{2.5}$$

$$h(b, \Phi) = \mathbb{E}[L(b(\omega), \omega)\Phi(\omega)]. \tag{2.6}$$

Define a function $\overline{L}$ on $\mathbb{R}^d$ by

$$\overline{L}(q) = \inf_{\substack{b:(b,\Phi)\in\mathcal{E} \\ \mathbb{E}(b\Phi)=q}} h(b,\Phi). \tag{2.7}$$

Observe that $\overline{L}$ is convex. This is a simple consequence of (2.7) and the fact that if $(b_i, \Phi_i) \in \mathcal{E}$, $i = 1, 2$, then for every $\lambda \in [0, 1]$,

$$\left(\frac{\lambda b_1 \Phi_1 + (1-\lambda) b_2 \Phi_2}{\lambda \Phi_1 + (1-\lambda)\Phi_2}, \lambda \Phi_1 + (1-\lambda)\Phi_2 \right) \in \mathcal{E}.$$

Let $\overline{H}$ be the convex conjugate of $\overline{L}$:

$$\begin{aligned} \overline{H}(p) &= \sup_{q\in\mathbb{R}^d} [\langle p, q\rangle - \overline{L}(q)] \\ &= \sup_{(b,\Phi)\in\mathcal{E}} [\langle p, m(b,\Phi)\rangle - h(b,\Phi)]. \end{aligned} \tag{2.8}$$

The solution $u(t,x)$ of (1.3), with $u(T,x) = U(x)$, is given by the Hopf-Lax-Oleĭnik formula

$$\begin{aligned} u(t,x) &= \sup_{q\in\mathbb{R}^d} \left[U(q) - (T-t)\overline{L}\left(\frac{q-x}{T-t}\right)\right] \\ &= \sup_{(b,\Phi)\in\mathcal{E}} [U(x + (T-t)m(b,\Phi)) - (T-t)h(b,\Phi)]. \end{aligned} \tag{2.9}$$

The main result of this paper is the following theorem:

THEOREM 2.2 *Assume that $H(p,\omega)$ satisfies* (H1)–(H3) *and that the terminal value $U(x)$ satisfies* (2.2) *on $\mathbb{R}^d$. Let $u(t,x)$ be the unique solution of* (1.2)–(1.3) *with $\overline{H}$ given by* (2.8). *Then with probability* 1, *for every $l > 0$,*

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

Remark 2.3. To identify the effective Hamiltonian it is enough to consider the set of linear terminal data $\{\langle p, x\rangle, \ p \in \mathbb{R}^d\}$. For each $p \in \mathbb{R}^d$ the terminal value problem for the averaged equation (1.3) with $U(x) = \langle p, x\rangle$ has an obvious solution $u(t,x) = \langle p,x\rangle - (T-t)\overline{H}(p)$. In particular, if we let $(t,x) = (0,0)$ and $T = 1$, then we get $\overline{H}(p) = u(0,0)$. Therefore if the homogenization result holds, then

$$\overline{H}(p) = \lim_{\varepsilon\to 0} u_\varepsilon^p(0,0,\omega) \quad \text{a.s.},$$

where u_ε^p solves (1.1)–(1.2) with $U(x) = \langle p, x\rangle$.

2.2 Shifts, Rescaling, and Variational Formulae

Our assumptions on the Hamiltonian allow us to use a variational representation of $u_\varepsilon(t,x,\omega)$, which can be considered the starting point of our analysis.

Denote by $\mathcal{C}$ the space of essentially bounded controls $c = c(s,x)$. Consider the diffusion on $\mathbb{R}^d$, with generator

$$\mathbf{A}_s = \frac{1}{2}\Delta + c(s,x)\cdot\nabla.$$

For $t < T$ and $x \in \mathbb{R}^d$, let $Q^c_{t,x}$ be the measure on $C([t,T);\mathbb{R}^d)$ that corresponds to the diffusion process with the above generator that starts from x at time t. For each $c \in \mathcal{C}$ and $\omega \in \Omega$ we set

$$v^c(t,x,\omega) = E^{Q^c_{t,x}}\left[U(x(T)) - \int_t^T L(c(s,x(s)),\tau_{(s,x(s))}\omega)\,ds\right].$$

Then

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

is the unique solution on $(-\infty,T]\times\mathbb{R}^d$ of

$$\frac{\partial v}{\partial t} + \frac{1}{2}\Delta v + H(\nabla v,\tau_{(t,x)}\omega) = 0 \tag{2.10}$$

with $v(T,x) = U(x)$ (see, for example, [11]).

An easy consequence of uniqueness is the covariant nature of the solution. If $v(\cdot,\cdot,\omega)$ is a solution corresponding to ω on $\mathbb{R}^d\times(-\infty,T]$, then $v^{s,y}(\cdot,\cdot,\omega) = v(\cdot+s,\cdot+y,\omega)$ is a solution on $(-\infty,T-s]\times\mathbb{R}^d$ corresponding to $\omega' = \tau_{(s,y)}\omega$ with terminal data $v^{s,y}(T-s,\cdot,\omega) = U(\cdot+y)$. In particular,

$$v(t,x,\omega) = v^{t,x}(0,0,\tau_{(t,x)}\omega) \tag{2.11}$$

if we match the terminal time and data.

A simple calculation shows $u_\varepsilon(t,x,\omega) = \varepsilon v_\varepsilon(t/\varepsilon,x/\varepsilon,\omega)$, where v_ε solves the unscaled equation (2.10) with the terminal condition

$$v_\varepsilon(T/\varepsilon,x,\omega) = U(\varepsilon x)/\varepsilon.$$

This leads to the following variational formula for the solution to the terminal value problem (1.1)–(1.2):

$$\begin{aligned} &u_\varepsilon(t,x,\omega) = \\ &\quad \sup_{c\in\mathcal{C}} E^{Q^c_{t/\varepsilon,x/\varepsilon}}\left[U(\varepsilon x(T/\varepsilon)) - \varepsilon\int_{t/\varepsilon}^{T/\varepsilon} L(c(s,x(s)),\tau_{(s,x(s))}\omega)\,ds\right]. \end{aligned} \tag{2.12}$$

It is sometimes more convenient to rescale the diffusion by considering the process $y_\varepsilon(s) = \varepsilon x(s/\varepsilon)$, $s\in[t,T]$, with the drift $c_\varepsilon(s,x) = c(s/\varepsilon,x/\varepsilon)\in\mathcal{C}$. Below we drop the subscript ε, since we take a supremum over all drifts in $\mathcal{C}$. After rescaling the generators become

$$\mathbf{A}_s^\varepsilon = \frac{\varepsilon}{2}\Delta + c(s,x)\cdot\nabla \tag{2.13}$$

with controls $c \in \mathcal{C}$. Writing $Q_{t,x}^{\varepsilon,c}$ for the measure associated with $\mathbf{A}_s^\varepsilon$, we get

$$\begin{aligned} u_\varepsilon(t,x,\omega) &= \\ &\sup_{c\in\mathcal{C}} E^{Q_{t,x}^{\varepsilon,c}}\left(U(y(T)) - \int_t^T L(c(s,y(s)), \tau_{(s/\varepsilon, y(s)/\varepsilon)}\omega)ds\right). \end{aligned} \tag{2.14}$$

Our goal is to show, locally uniform in t and x, the almost sure convergence of $u_\varepsilon(t,x,\omega)$ to $u(t,x)$ given by (2.9).

2.3 Ergodic Theorem and a Lower Bound

Let $b \in \mathbf{B} \equiv L^\infty(\Omega;\mathbb{R}^d)$. Consider a Brownian motion $x(t)$ on $\mathbb{R}^d$ starting from 0 at time 0 with a random drift $b(\tau_{(t,x)}\omega)$, and denote the corresponding measure on $C([0,\infty);\mathbb{R}^d)$ by $Q_{0,0}^{b,\omega}$. Then this diffusion can be "lifted" to Ω as follows. Pick a starting point ω. Define a path $\omega(s)$, which starts from ω at time 0, by $\omega(s) = \tau_{(s,x(s))}\omega$, $s \geqslant 0$.

The measure $P^{b,\omega}$ induced on paths in Ω corresponds to a Markov process on Ω with the generator

$$\mathcal{A}_b = D_t + \frac{1}{2}\Delta + b(\omega)\cdot\nabla.$$

If we can find a positive density Φ on Ω such that $\Phi\, d\mathbb{P}$ is an invariant ergodic probability measure for $\mathcal{A}_b$, then by the ergodic theorem for every $F \in L^1(\Omega, \Phi\, d\mathbb{P})$

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

$P^{b,\omega}$-a.s. or in $L^1(P^{b,\omega})$ for $\mathbb{P}$-a.e. ω.

Finding an invariant density for a given drift is a hard problem. We work in Ω, which is infinite dimensional, while our diffusion is finite dimensional and is very degenerate when viewed on Ω. It is clear, though, that $\mathcal{E}$ is not empty, since it obviously contains pairs $(b, 1)$, where b is a constant. Moreover, it is easy to show that if $(b,\Phi) \in \mathcal{E}$, then $\Phi\, d\mathbb{P}$ is an *ergodic* invariant measure for $\mathcal{A}_b$, and we can use the ergodic theorem.

If we view ω in the formula (2.12) for u_ε as a parameter, we may allow the controls c to be dependent on ω as well. Let us consider only stationary controls, i.e., $c(t,x,\omega) = b(\tau_{(t,x)}\omega)$, where b is such that $(b,\Phi) \in \mathcal{E}$ for some Φ. Setting $(t,x) = 0$, for a.e. ω with respect to $\mathbb{P}$, we obtain almost surely with respect to

$P^{b,\omega}$ and in $L^1(P^{b,\omega})$

$$\lim_{\varepsilon\to 0} \varepsilon \int_0^{T/\varepsilon} b(\omega(s))ds = T \int_\Omega b(\omega)\Phi(\omega)d\,\mathbb{P} = Tm(b,\Phi),$$

$$\lim_{\varepsilon\to 0} \varepsilon \int_0^{T/\varepsilon} L(b(\omega(s)),\omega(s))ds = T \int_\Omega L(b(\omega),\omega)\Phi\, d\,\mathbb{P} = Th(b,\Phi).$$

Therefore, for the diffusions $Q_{0,0}^{b,\omega}$ on $\mathbb{R}^d$ we obtain, for almost all ω with respect to $\mathbb{P}$,

$$\lim_{\varepsilon\to 0} E^{Q_{0,0}^{b,\omega}}[\varepsilon|x(T/\varepsilon) - Tm(b,\Phi)|] = 0, \tag{2.15}$$

$$\lim_{\varepsilon\to 0} E^{Q_{0,0}^{b,\omega}}\left[\varepsilon\left|\int_0^{T/\varepsilon} L(b(\tau_{(t,x(t))}\omega,\omega)dt - Th(b,\Phi)\right|\right] = 0. \tag{2.16}$$

It now follows that for each $(b,\Phi)\in\mathcal{E}$,

$$\liminf_{\varepsilon\to 0} u_\varepsilon(0,0,\omega) \geqslant U(Tm(b,\Phi)) - Th(b,\Phi) \quad \mathbb{P}\text{-a.s.} \tag{2.17}$$

Notice that, because of (2.9), this establishes an almost sure lower bound at $(0,0)$. The relation (2.11) and the translation invariance of $\mathbb{P}$ imply that the lower bound holds for arbitrary (t,x) but in probability. More work needs to be done to obtain an almost sure locally uniform lower bound. The proof is similar to the one given in Section 4 of [14], and the details are presented in the next section.

Remark (A few words about an upper bound). An upper bound is essentially obtained by comparison with a family of supersolutions of (1.1). The starting point of the construction is formula (2.8) for the effective Hamiltonian. The main idea is the same as in [14]. There are some difficulties in the construction due to the lack of control on the time derivatives. This problem was not present in the case of the time-independent Hamiltonian.

3 Lower Bound

We begin with an easy estimate on the solution $u_\varepsilon(t,x,\omega)$ of (1.1) and its approximations

$$u_\varepsilon^k(t,x,\omega) \overset{\text{def}}{=} \sup_{c\in\mathcal{C}_k} E^{Q_{t,x}^{\varepsilon,c}}\Big(U(y(T)) - \int_t^T L(c(s,y(s)),\tau_{(s/\varepsilon,y(s)/\varepsilon)}\omega)ds\Big), \tag{3.1}$$

where $y(\cdot)$ is a diffusion with the generator (2.13) and $\mathcal{C}_k$ consists of $c\in\mathcal{C}$ that are bounded in norm by k.

LEMMA 3.1 *Assume* (L1) *and* (2.2). *Then there are positive constants* C_1 *and* C_2 *such that*

$$|u_\varepsilon(t,x,\omega) - U(x)| \leqslant C_1(T-t)^{1/2} + C_2|T-t|,$$
$$|u_\varepsilon^k(t,x,\omega) - U(x)| \leqslant C_1(T-t)^{1/2} + C_2|T-t|.$$

PROOF: From the variational formula (2.14), the bounds (L1), and the convexity of $|q|^\alpha$ for $\alpha \geqslant 1$, it follows that, uniformly in ω,

$$\begin{aligned} u_\varepsilon(t,x,\omega) - U(x) &\leqslant \sup_{c\in\mathcal{C}} \bigg[K\bigg[E^{Q_{t,x}^{\varepsilon,c}} \int_t^T |c(s,x(s))|ds \bigg] + CK\varepsilon^{1/2}(T-t)^{1/2} \\ &\qquad - E^{Q_{t,x}^{\varepsilon,c}} \int_t^T (c_4|c(s,x(s))|^{\beta'} - c_5)ds \bigg] \\ &\leqslant CK\varepsilon^{1/2}(T-t)^{1/2} + (T-t) \sup_{q\in\mathbb{R}^d} [K|q| - c_4|q|^{\beta'} + c_5]. \end{aligned}$$

For the lower bound we can take $c = 0$ to get

$$u_\varepsilon(t,x,\omega) - U(x) \geqslant -C\varepsilon^{1/2}(T-t)^{1/2} - (T-t)\sup_\omega L(0,\omega).$$

The same proof works for u_ε^k as well. □

In the previous section we established a weak version of the lower bound. We state it now as a lemma.

LEMMA 3.2 *Assuming the regularity condition* (2.2) *on* U, *for almost all* ω *with respect to* $\mathbb{P}$,

$$\liminf_{\varepsilon\to 0} [u_\varepsilon(0,0,\omega) - u(0,0)] \geqslant 0. \tag{3.2}$$

PROOF: Relation (2.17) states that for each $(b,\Phi) \in \mathcal{E}$

$$\liminf_{\varepsilon\to 0} [u_\varepsilon(0,0,\omega) - U(Tm(b,\varphi)) + Th(b,\Phi)] \geqslant 0,$$

$\mathbb{P}$ almost surely. Since the supremum over $\mathcal{E}$ is easily reduced to the supremum over a countable subset of $\mathcal{E}$ in the formula (2.9), the assertion (3.2) of the lemma follows. □

The next lemma strengthens the estimate (3.2).

LEMMA 3.3 *Under the conditions of Lemma* 3.2, *for each* $\eta > 0$, *there is a subset* $N_\eta \subset \Omega$ *such that* $\mathbb{P}(N_\eta) \geqslant 1 - \eta$ *and*

$$\liminf_{r\to 0} \liminf_{\varepsilon\to 0} \inf_{\omega\in N_\eta} \inf_{\substack{|y|\leqslant r \\ |t|\leqslant r}} [u_\varepsilon(t,y,\omega) - u(0,0)] \geqslant 0. \tag{3.3}$$

PROOF: It is enough to prove that for almost all ω with respect to $\mathbb{P}$,

$$\liminf_{r\to 0}\liminf_{\varepsilon\to 0}\inf_{\substack{|y|\leqslant r\\|t|\leqslant r}}[u_\varepsilon(t,y,\omega)-u(0,0)]\geqslant 0, \tag{3.4}$$

and (3.3) will follow by an application of Egoroff's theorem.

Consider functions u_ε^k defined by (3.1). Clearly, $u_\varepsilon^k(t,x,\omega)\nearrow u_\varepsilon(t,x,\omega)$ as $k\uparrow\infty$. Therefore, by Lemma 3.2 it is sufficient to prove that for each $k<\infty$,

$$\liminf_{r\to 0}\liminf_{\varepsilon\to 0}\inf_{\omega}\inf_{\substack{|y|\leqslant r\\|t|\leqslant r}}[u_\varepsilon(t,y,\omega)-u_\varepsilon^k(0,0,\omega)]\geqslant 0. \tag{3.5}$$

We break the proof into two parts. For fixed k,

$$\liminf_{r\to 0}\liminf_{\varepsilon\to 0}\inf_{\omega}\inf_{|y|\leqslant r}[u_\varepsilon^{k+1}(0,y,\omega)-u_\varepsilon^k(0,0,\omega)]\geqslant 0, \tag{3.6}$$

$$\liminf_{r\to 0}\liminf_{\varepsilon\to 0}\inf_{\omega}\inf_{\substack{|y|\leqslant r\\|t|\leqslant r}}[u_\varepsilon^{k+1}(t,y,\omega)-u_\varepsilon^k(0,y,\omega)]\geqslant 0. \tag{3.7}$$

Step 1. To prove (3.6) we let $x(\cdot)$ be a path of a diffusion with a drift $b\in\mathcal{C}_k$ starting from 0 at time 0. We shall couple it to $y(\cdot)$, which starts from y at time 0 as follows: Assume that $r<T$. Define

$$y(s)=\begin{cases} y+x(s)-\frac{sy}{|y|}, & 0\leqslant s<|y|,\\ x(s), & |y|\leqslant s\leqslant T.\end{cases}$$

Then $y(\cdot)$ has the drift

$$c(s,z)=\begin{cases} b\big(s,z-y+\frac{sy}{|y|}\big)-\frac{y}{|y|} & \text{if } 0\leqslant s<|y|,\\ b(s,z) & \text{if } |y|\leqslant s\leqslant T,\end{cases}$$

and $c\in\mathcal{C}_{k+1}$. Let us denote the coupled measure by $Q^{\varepsilon,b,c}_{0,0,y}$.

Since the two paths are identical after time $|y|$, they end up at the same place (if $|y|\leqslant r$) and

$$\begin{aligned} &u_\varepsilon^k(0,0,\omega)-u_\varepsilon^{k+1}(0,y,\omega)\\ &\quad\leqslant\sup_{b\in\mathcal{C}_k}\bigg[E^{Q^{\varepsilon,b,c}_{0,0,y}}\int_0^{|y|}|L(b(s,x(s)),\tau_{(s/\varepsilon,x(s)/\varepsilon)}\omega)|\,ds\\ &\qquad\qquad+E^{Q^{\varepsilon,b,c}_{0,0,y}}\int_0^{|y|}|L(c(s,y(s)),\tau_{(s/\varepsilon,y(s)/\varepsilon)}\omega)|\,ds\bigg]\\ &\quad\leqslant 2|y|\sup_{\substack{\omega\\|q|\leqslant k+1}}|L(q,\omega)|.\end{aligned}$$

This proves (3.6).

Remark 3.4. Observe that by using the same argument as in Step 1 we can show that for a fixed k

$$\liminf_{r\to 0}\liminf_{\varepsilon\to 0}\inf_{\omega}\inf_{t\leqslant T}\inf_{|y-z|\leqslant r}[u_\varepsilon^{k+1}(t,y,\omega)-u_\varepsilon^k(t,z,\omega)]\geqslant 0. \tag{3.8}$$

Step 2. Now we turn to (3.7). Let us take the case $t<0$. Then we can choose $c(s,x)=0$ until time 0 and then make the optimal choice from $\mathcal{C}_{k+1}$. Therefore

$$u_\varepsilon^{k+1}(t,y,\omega)\geqslant\int_{\mathbb{R}^d}u_\varepsilon^{k+1}(0,z,\omega)p_\varepsilon(t,y;0,z)dz-|t|\sup_{\omega}L(0,\omega) \tag{3.9}$$

where p_ε is the fundamental solution of $\frac{\partial}{\partial t}+\frac{\varepsilon}{2}\Delta$. From Remark 3.4 we can obtain a uniform lower bound on $u_\varepsilon^{k+1}(0,z,\omega)-u_\varepsilon^k(0,y,\omega)$ provided $|z|$ is small. In addition, we have for fixed $r'>0$,

$$\lim_{r\to 0}\limsup_{\varepsilon\to 0}\sup_{\substack{|y|\leqslant r\\|t|\leqslant r}}\int_{|z|>r'}p_\varepsilon(t,y;0,z)dz=0.$$

From the bounds on u_ε^{k+1} of Lemma 3.1 and from (2.2), it is clear that

$$\lim_{r\to 0}\limsup_{\varepsilon\to 0}\sup_{\substack{|y|\leqslant r\\|t|\leqslant r}}\int_{|z|>r'}|u_\varepsilon^{k+1}(0,z,\omega)|p_\varepsilon(t,y;0,z)dz=0,$$

which together with (3.9) implies (3.7).

The case when $t>0$ is similar. We represent $u_\varepsilon^k(0,y,\omega)$ by a variational formula, make the optimal choice of $c\in\mathcal{C}_k$, and get

$$u_\varepsilon^{k+1}(t,y,\omega)-u_\varepsilon^k(0,y,\omega)\geqslant$$
$$\int_{\mathbb{R}^d}\big(u_\varepsilon^{k+1}(t,y,\omega)-u_\varepsilon^k(t,z,\omega)\big)p_\varepsilon^c(0,y;t,z)dz-t\sup_{|q|\leqslant k}|L(q,\omega)|,$$

where p_ε^c is the fundamental solution of $\frac{\partial}{\partial t}+c(t,x)\cdot\nabla+\frac{\varepsilon}{2}\Delta$. Using (3.8), Lemma 3.1, condition (2.2), and properties of p_ε^c with $c\in\mathcal{C}_k$, we can derive (3.7). □

Combining the relation (2.11), Lemma 3.3, and the ergodic theorem, we obtain the almost sure locally uniform lower bound.

THEOREM 3.5 *Assume that* $H(p,\omega)$ *satisfies* (H1) *and* (H2) *and that the terminal condition* $U(x)$ *is uniformly continuous on* $\mathbb{R}^d$. *Let* $u(t,x)$ *be the unique solution of* (1.2)–(1.3) *with* $\overline{H}$ *given by* (2.8). *Then with probability* 1, *for every* $l>0$,

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

PROOF: We omit the proof, since it is essentially the same as that for theorem 2.1 in [14]. □

4 Construction of Approximate Supersolutions

Let $H(p,\omega)$ satisfy (H1)–(H3). Fix an arbitrary $\theta \in \mathbb{R}^d$. Our goal is ideally to construct, for each $\theta \in \mathbb{R}^d$, a supersolution V such that

$$\partial_t V + \frac{1}{2}\Delta V + H(\theta + \nabla V, \tau_{(t,x)}\omega) \leqslant \overline{H}(\theta), \quad (t,x) \in \mathbb{R}^{d+1}.$$

Actually, it is enough to construct approximate supersolutions $V_\delta(t,x,\omega)$, $\delta \in (0,\delta_0)$, or even just their space-time gradients with mean 0, such that, for a.e. ω,

$$\partial_t V_\delta + \frac{1}{2}\Delta V_\delta + H(\theta + \nabla V_\delta, \tau_{(t,x)}\omega) \leqslant \overline{H}(\theta) + \nu(\delta), \quad (t,x) \in \mathbb{R}^{d+1}, \tag{4.1}$$

where $\nu(\delta) > 0$, $\nu(\delta) \to 0$ as $\delta \to 0$.

The supersolution V is constructed by duality very much like in the time-independent case. At a purely formal level there is no difference. However, in terms of the regularity of the supersolution there is considerable difference. One can only establish the minimal regularity for the supersolution. The inequality (4.1) and the lower bound of (H1) allow us to control $E[|\nabla V|^\alpha]$. The bounds one gets are local L^α bounds on ∇V and an L^∞ upper bound on $\partial_t V + \frac{1}{2}\Delta V$. The fact that $\partial_t V + \frac{1}{2}\Delta V$ has mean 0 provides a local L^1 bound on $\partial_t V + \frac{1}{2}\Delta V$. Convolution in space by a smoothing kernel will get an estimate on ΔV, and this will finally allow us to control $V_t * \varphi$ in L^1_{loc} for smooth φ. For this one has to work with weak compactness in L^1, which is always delicate.

An upper bound on u_ε will then be established in the next section using a probabilistic comparison argument and the ergodic theorem, Theorem 5.3.

At first, we introduce some notation. Let us recall that $\mathbf{B} \equiv L^\infty(\Omega;\mathbb{R}^d)$. We define $\mathbf{B}_r = \{b \in \mathbf{B} : \|b\|_\infty \leqslant r\}$ and

$$H_r(p,\omega) = \sup_{|q|\leqslant r} (\langle p,q\rangle - L(q,\omega)).$$

Clearly $\mathbf{B}_r$ is compact and convex in $L^p(\Omega,\mathbb{P})$ in the weak topology for $1 \leqslant p < \infty$. For each $k > 1$ we define

$$\mathbf{E}_k = \left\{\Phi : k^{-1} \leqslant \Phi(\omega) \leqslant k,\ \int \Phi(\omega)\,d\mathbb{P} = 1\right\},$$

$$\mathbf{D}_k = \left\{\Phi : k^{-1} \leqslant \Phi(\omega) \leqslant k,\ |\nabla\Phi| \leqslant k^2,\ \int \Phi(\omega)\,d\mathbb{P} = 1\right\} \subset \mathbf{E}_k,$$

which are again compact convex sets if appropriate weak topologies are imposed. Observe that $\mathcal{E} = \bigcup_{k>0}\bigcup_{r>0}\mathcal{E}_{r,k}$, where

$$\mathcal{E}_{r,k} = \{(b,\Phi) \in \mathbf{B}_r \times \mathbf{D}_k : D_t\Phi + \operatorname{div}(b\Phi) = \tfrac{1}{2}\Delta\Phi\}.$$

If φ is a mollifier on $\mathbb{R}^d$ and $F \in L^1(\Omega,\mathbb{P})$, we can define

$$F^\varphi(\omega) = \int F(\tau_{(0,x)}\omega)\varphi(x)\,dx. \tag{4.2}$$

Then $F^{\varphi} \in L^1(\Omega, \mathbb{P})$ and is smooth under space shifts. Notice that if $\Phi \in \mathbf{E}_k$, then $\Phi^{\varphi}(\omega) \in \mathbf{D}_k$ provided k is large enough so that $\|\nabla\varphi\|_\infty \leqslant k$.

We start with formula (2.8) for the effective Hamiltonian and use the minimax theorem to obtain a dual formula.

LEMMA 4.1 *There is a sequence of functions* $\{F_n(\omega)\}$, $n \geqslant 1$, $F_n \in W^{1,\infty}(\Omega)$, $F_n(\tau_{(\cdot,\cdot)}\omega) \in C^\infty(\mathbb{R}^{d+1})$ *a.s. such that*

$$\sup_{\Phi\in\mathbf{D}_n} \mathbb{E}[[D_t F_n + \tfrac{1}{2}\Delta F_n + H_n(\theta + \nabla F_n, \omega)]\Phi] \leqslant \overline{H}(\theta) + \tfrac{1}{n}. \tag{4.3}$$

PROOF: From (2.8) we have

$$\overline{H}(\theta) \geqslant \sup_{(b,\Phi)\in\mathcal{E}_{r,k}} \mathbb{E}[(\langle\theta, b(\omega)\rangle - L(b(\omega),\omega))\Phi(\omega)].$$

The definition of $\mathcal{E}_{r,k}$ contains a constraint. Let $\mathcal{Y}$ be a space of nice test functions (for example, $\mathcal{Y}$ could be a class of functions $F \in W^{1,\infty}(\Omega)$ convoluted with a mollifier from $C_0(\mathbb{R}^{d+1})$ to ensure the smoothness under the shifts). Using the fact that for $\Phi \in \mathbf{D}_k$

$$\begin{aligned}
&\inf_{F\in\mathcal{Y}} \mathbb{E}([D_t F + \langle b, \nabla F\rangle + \tfrac{1}{2}\Delta F]\Phi) \\
&\quad = \inf_{F\in\mathcal{Y}} \mathbb{E}([D_t F + \langle b, \nabla F\rangle]\Phi - \tfrac{1}{2}\langle\nabla F, \nabla\Phi\rangle) \\
&\quad = \begin{cases} 0 & \text{if } D_t\Phi + \operatorname{div}(b\Phi) = \tfrac{1}{2}\Delta\Phi \text{ in the weak sense,} \\ -\infty & \text{otherwise,} \end{cases}
\end{aligned}$$

we can remove the constraint and get the inequality

$$\begin{aligned}
\overline{H}(\theta) &\geqslant \sup_{\Phi\in\mathbf{D}_k} \sup_{b\in\mathbf{B}_r} \inf_{F\in\mathcal{Y}} \mathbb{E}[(D_t F + \langle b, \theta + \nabla F\rangle - L(b,\omega) + \tfrac{1}{2}\Delta F)\Phi] \\
&= \sup_{\Phi\in\mathbf{D}_k} \sup_{b\in\mathbf{B}_r} \inf_{F\in\mathcal{Y}} \mathbb{E}[(D_t F + \langle b, \theta + \nabla F\rangle - L(b,\omega))\Phi - \tfrac{1}{2}\langle\nabla F, \nabla\Phi\rangle].
\end{aligned}$$

The set $\mathbf{B}_r$ is convex and compact in the weak topology while $\mathcal{Y}$ is convex. For fixed Φ, the functional

$$\mathbb{E}[(D_t F + \langle b, \theta + \nabla F\rangle - L(b,\omega))\Phi - \tfrac{1}{2}\langle\nabla F, \nabla\Phi\rangle]$$

is concave and upper semicontinuous in b in the weak topology and is linear in F. We can therefore apply Sion's minimax theorem [20] and interchange the infimum and the supremum to obtain

$$\overline{H}(\theta) \geqslant \sup_{\Phi\in\mathcal{D}_k} \inf_{F\in\mathcal{Y}} \sup_{b\in\mathbf{B}_r} \mathbb{E}[(D_t F + \langle b, \theta + \nabla F\rangle - L(b,\omega))\Phi - \tfrac{1}{2}\langle\nabla F, \nabla\Phi\rangle].$$

For fixed Φ and F the supremum over $b \in \mathbf{B}_r$ can be taken inside the expectation. Recalling the definition of $H_r(p,\omega)$ we get

$$\overline{H}(\theta) \geqslant \sup_{\Phi\in\mathbf{D}_k} \inf_{F\in\mathcal{Y}} \mathbb{E}[(D_t F + H_r(\theta + \nabla F, \omega))\Phi - \tfrac{1}{2}\langle\nabla F, \nabla\Phi\rangle].$$

Now we would like to use the minimax theorem one more time. The functional

$$\mathbb{E}[(D_t F + H_r(\theta + \nabla F, \omega))\Phi - \tfrac{1}{2}\langle \nabla F, \nabla \Phi \rangle]$$

is convex in F and is a continuous linear functional of Φ in the weak topology (requiring the weak convergence of Φ and $\nabla\Phi$). Therefore, for each k and r

$$\overline{H}(\theta) \geqslant \inf_{F \in \mathcal{Y}} \sup_{\Phi \in \mathbf{D}_k} \mathbb{E}[(D_t F + \tfrac{1}{2}\Delta F + H_r(\theta + \nabla F, \omega))\Phi].$$

This immediately implies the statement of the lemma. □

Let us briefly describe the idea of the next step. We would like to take some sort of weak limit of F_n and produce an F such that for almost all ω with respect to $\mathbb{P}$,

$$D_t F(\omega) + \frac{1}{2}\operatorname{div}(\nabla F(\omega)) + H(\theta + \nabla F(\omega), \omega) \leqslant \overline{H}(\theta).$$

Unfortunately, F may not exist. Let us first see what estimates one might expect on F. Taking expectations, one can see that $\mathbb{E}[H(\theta + \nabla F(\omega), \omega)]$ and therefore $\mathbb{E}[|\nabla F|^\alpha]$ can be controlled. With some spatial (in x) mollification, for the convolution $F^\varphi = F * \varphi$, one can get an upper bound on $\mathbb{E}[(D_t F^\varphi)^+]$. Since $D_t F$ should have mean 0, this means one can expect to control $\mathbb{E}[|D_t F^\varphi|]$. At best, we will be able to produce a limiting family of objects f_φ, g such that $\mathbb{P}$-a.s.

$$f_\varphi(\omega) + \frac{1}{2}\operatorname{div} g^\varphi(\omega) + \int H(\theta + g(\tau_{(0,x)}\omega), \tau_{(0,x)}\omega)\varphi(x)dx \leqslant \overline{H}(\theta), \tag{4.4}$$

where

$$g^\varphi = \int g(\tau_{(0,x)}\omega)\varphi(x)dx, \quad \operatorname{div} g^\varphi = -\int \langle g(\tau_{(0,x)}\omega), \nabla\varphi(x)\rangle dx. \tag{4.5}$$

We shall end up with $g \in L^\alpha(\Omega, \mathbb{P})$ and a collection $f_\varphi \in L^1(\Omega, \mathbb{P})$, not necessarily smooth with respect to the shifts, which satisfies for almost all ω and each φ the compatibility condition

$$\int f_\varphi(\tau_{(t,x)}\omega)(\operatorname{div} G(t,x))dt\,dx = \int \langle g^\varphi(\tau_{(t,x)}\omega), G_t(t,x)\rangle dt\,dx \tag{4.6}$$

for $G \in C_0^\infty(\mathbb{R} \times \mathbb{R}^d; \mathbb{R}^d)$. We shall lose control of $\|f_\varphi\|_1$ as $\varphi \to \delta_0$ in $\mathcal{D}'(\mathbb{R}^d)$.

The details of the above construction are presented in the next theorem.

THEOREM 4.2 *Let H satisfy* (H1) *and* (H2). *For every $\theta \in \mathbb{R}^d$ there is $g : \Omega \to \mathbb{R}^d$, $g \in L^\alpha(\Omega, \mathbb{P})$, and for every $r > 0$ there is a nonnegative mollifier $\varphi \in C_0^\infty(\mathbb{R}^d)$, $\varphi(x) = 0$ outside of $B(0,r)$, and a function $f_\varphi \in L^1(\Omega, \mathbb{P})$ such that*

(i) $\mathbb{E}[f_\varphi] = 0$, $\mathbb{E}[g] = \theta$;
(ii) $\nabla \times g^\varphi = 0$ *and* (4.6) *holds, where g^φ is defined in* (4.5);
(iii) *inequality* (4.4) *holds $\mathbb{P}$-a.s.*;

(iv) *there is a constant* $K(\varphi) > 0$, *which depends only on* φ, c_1, α, *and* $\overline{H}(\theta)$, *such that*

$$f_\varphi(\omega) + \frac{c_1}{2}|g^\varphi(\omega)|^\alpha \leqslant K(\varphi) \quad \mathbb{P}\text{-}a.s.$$

PROOF: Let $f_n = D_t F_n$ and $g_n = \theta + \nabla F_n$, where $\{F_n\}$, $n \geqslant 1$, is constructed in Lemma 4.1.

Step 1. The sequence $\{g_n\}$ is uniformly integrable (see theorem 5.2 of [14]). Take a weakly convergent subsequence and call it again $\{g_n\}$. We have $g_n \to g$ as $n \to \infty$ weakly in $L^1(\Omega, \mathbb{P})$ and

$$\mathbb{E}[H_n(g_n(\omega), \omega)] \leqslant \overline{H}(\theta) + \frac{1}{n}.$$

Clearly, $H_k \leqslant H_n$ for $k \leqslant n$. Then

$$\mathbb{E}[H_k(g_n(\omega), \omega)] \leqslant \overline{H}(\theta) + \frac{1}{n}.$$

Let $n \to \infty$ and use the weak convergence of g_n to g and the convexity of H_k to get $\mathbb{E}[H_k(g(\omega), \omega)] \leqslant \overline{H}(\theta)$. Finally, we let $k \to \infty$ to obtain

$$\mathbb{E}[H(g(\omega), \omega)] \leqslant \overline{H}(\theta).$$

Using the lower bound on H, we conclude that $g \in L^\alpha(\Omega, \mathbb{P})$.

Step 2. The only control we have on $\{f_n\}$, $n \geqslant 1$, is inequality (4.3) of Lemma 4.1:

$$\sup_{\Phi \in \mathbf{D}_n} \mathbb{E}\left[\left(f_n + \frac{1}{2}\operatorname{div} g_n + H_n(g_n, \omega)\right)\Phi\right] \leqslant \overline{H}(\theta) + \frac{1}{n}. \tag{4.7}$$

This control is very weak. We mollify f_n by convoluting with φ and obtain some estimates on $f_n^\varphi = f_n * \varphi$, $n \geqslant 1$.

Take a mollifier $\varphi \in C_0^\infty(\mathbb{R}^d)$, $\varphi(x) \geqslant 0$, and $\int \varphi(x)dx = 1$ such that

$$\int |\nabla\varphi(x)|dx \leqslant \left(\int \left|\frac{\nabla\varphi(x)}{\varphi(x)}\right|^{\alpha'} \varphi(x)dx\right)^{1/\alpha'} = C(\varphi) < \infty \tag{4.8}$$

where $(\alpha')^{-1} + \alpha^{-1} = 1$ and $C(\varphi)$ is a constant that depends on φ. We will continue to denote any constant that depends only on φ by $C(\varphi)$, even if it changes during a step. Our first goal is to establish a uniform bound on $\mathbb{E}[|f_n^\varphi|]$. We rewrite (4.7) in the form

$$\mathbb{E}[f_n\Phi] \leqslant -\mathbb{E}\left[\left(\frac{1}{2}\operatorname{div} g_n + H_n(g_n, \omega)\right)\Phi\right] + \overline{H}(\theta) + \frac{1}{n}$$

valid for all $\Phi \in \mathbf{D}_n$, or

$$\mathbb{E}[f_n\Phi^\varphi] \leqslant -\mathbb{E}\left[\left(\frac{1}{2}\operatorname{div} g_n + H_n(g_n, \omega)\right)\Phi^\varphi\right] + \overline{H}(\theta) + \frac{1}{n} \tag{4.9}$$

valid for $\Phi \in \mathbf{E}_n$. We note that H_n has a uniform lower bound $H_n \geqslant -C$. Moving the derivatives to φ as in (4.5), we get

$$\left|\mathbb{E}[\tfrac{1}{2}(\operatorname{div} g_n)\Phi^\varphi]\right| \leqslant C(\varphi)\mathbb{E}[|g_n|]\|\Phi\|_\infty \leqslant cC(\varphi)\|\Phi\|_\infty,$$

where c is a uniform bound on $\mathbb{E}[|g_n|]$. Therefore, for all $\Phi \in \mathbf{E}_n$ there are constants c, C, and $C(\varphi)$ such that

$$\mathbb{E}[f_n^\varphi \Phi] = \mathbb{E}[f\Phi_n^\varphi] \leqslant cC(\varphi)\|\Phi\|_\infty + C + \overline{H}(\theta).$$

If $\Phi > 0$ is not normalized but $\Phi/\mathbb{E}[\Phi] \in \mathbf{E}_n$, then

$$\mathbb{E}[f_n^\varphi \Phi] \leqslant cC(\varphi)\|\Phi\|_\infty + (C + \overline{H}(\theta))\mathbb{E}[\Phi].$$

Consider functions Φ_n, $n = 1, 2, \dots$, defined by

$$\Phi_n(\omega) = \begin{cases} \frac{1}{2} & \text{if } f_n^\varphi(\omega) \leqslant 0, \\ 2 & \text{if } f_n^\varphi(\omega) > 0. \end{cases}$$

Since $\mathbb{E}[f_n^\varphi] = 0$, we have

$$\mathbb{E}[f_n^\varphi : f_n^\varphi > 0] = -\mathbb{E}[f_n^\varphi : f_n^\varphi \leqslant 0] = \frac{1}{2}\,\mathbb{E}[|f_n^\varphi|].$$

Also, $\Phi_n/\mathbb{E}[\Phi_n] \in \mathbf{E}_4$, $\mathbb{E}[\Phi] \leqslant \|\Phi\|_\infty \leqslant 2$, and

$$\mathbb{E}[f_n^\varphi \Phi_n] = 2\,\mathbb{E}[f_n^\varphi : f_n^\varphi > 0] + \frac{1}{2}\,\mathbb{E}[f_n^\varphi : f_n^\varphi \leqslant 0] = \frac{3}{4}\,\mathbb{E}[|f_n^\varphi|]$$

giving us the bound

$$\sup_{n \geqslant n_0(\varphi)} \mathbb{E}[|f_n^\varphi|] \leqslant C(\theta, \varphi). \tag{4.10}$$

Step 3. We would like to prove the uniform integrability of $\{f_n^\varphi\}$, but at this point we can only establish the uniform integrability of the positive parts of $\{f_n^\varphi\}$, i.e., the uniform integrability of $\{f_n^{\varphi,+}\}$ where $f_n^{\varphi,+} = \max\{f_n^\varphi, 0\}$. This depends on the uniform integrability of g_n established in Step 1. If we truncate $|g_n|$ at level N and write $g_n = g_{n,N} + (g_n - g_{n,N})$, we have for all $\Phi \in \mathbf{E}_k$

$$-\frac{1}{2}\,\mathbb{E}\langle g_n, \nabla\Phi^\varphi\rangle \leqslant \mathbb{E}[|g_n - g_{n,N}|]\|\nabla\Phi^\varphi\|_\infty - \frac{1}{2}\,\mathbb{E}\langle g_{n,N}, \nabla\Phi^\varphi\rangle, \tag{4.11}$$

and $\sup_n \mathbb{E}[|g_n - g_{n,N}|] = \delta(N) \to 0$ as $N \to \infty$ by uniform integrability. Moreover,

$$-\frac{1}{2}\,\mathbb{E}\langle g_{n,N}, \nabla\Phi^\varphi\rangle = \frac{1}{2}\,\mathbb{E}[(\operatorname{div}\, g_{n,N}^\varphi)\Phi].$$

We can estimate

$$|\operatorname{div} g_{n,N}^\varphi| = \left|\int \langle g_{n,N}(\tau_{(0,x)}\omega), \nabla\varphi(x)\rangle\, dx\right|$$
$$\leqslant C(\varphi)\left(\int |g_{n,N}(\tau_{(0,x)}\omega)|^\alpha \varphi(x) dx\right)^{1/\alpha}.$$

On the other hand, for $n \geqslant n_0(N)$ large enough,

$$H_n(g_{n,N}) = H(g_{n,N}) \geqslant c_1 |g_{n,N}|^\alpha - c_2$$

and

$$\begin{aligned}\mathbb{E}[H_n(g_n(\omega),\omega))\Phi^\varphi(\omega)] &\geqslant \mathbb{E}[(c_1|g_{n,N}(\omega)|^\alpha - c_2)\Phi^\varphi(\omega)] \\ &= c_1 \mathbb{E}\left[\Phi(\omega)\int |g_{n,N}(\tau_{(0,x)}\omega)|^\alpha \varphi(x)\,dx\right] - c_2.\end{aligned}$$

Therefore,

$$\begin{aligned}&-\frac{1}{2}\mathbb{E}\langle g_{n,N}, \nabla\Phi^\varphi\rangle - \mathbb{E}[H_n(g_n(\omega),\omega))\Phi^\varphi(\omega)] \\ &\leqslant \frac{1}{2}C(\varphi)\mathbb{E}\left[\Phi(\omega)\left(\int |g_{n,N}(\tau_{(0,x)}\omega)|^\alpha\varphi(x)\,dx\right)^{1/\alpha}\right] \\ &\quad - c_1\mathbb{E}\left[\Phi(\omega)\int |g_{n,N}(\tau_{(0,x)}\omega)|^\alpha\varphi(x)\,dx\right] + c_2 \\ &\leqslant \max_{x\geqslant 0}\left(\frac{1}{2}C(\varphi)x - c_1 x^\alpha\right) + c_2 \leqslant C(\varphi).\end{aligned}$$

Combining the above inequality with (4.9) and (4.11), we conclude that for all $n \geqslant n_0(N)$

$$\mathbb{E}[f_n^\varphi \Phi] \leqslant \delta(N)\|\nabla\Phi^\varphi\|_\infty + C(\theta,\varphi).$$

Finally, letting $n \to \infty$ and then $N \to \infty$, we obtain for every k

$$\limsup_{n\to\infty}\sup_{\Phi\in\mathbf{E}_k}\mathbb{E}[f_n^\varphi\Phi] \leqslant C(\theta,\varphi). \tag{4.12}$$

Consider functions $\Phi_{n,k,\ell}$ defined by

$$\Phi_{n,k,\ell}(\omega) = \begin{cases}\frac{1}{k} & \text{if } f_n^\varphi(\omega) \leqslant \ell, \\ k & \text{if } f_n^\varphi(\omega) > \ell.\end{cases}$$

Then $\Phi_{n,k,\ell}/\mathbb{E}[\Phi_{n,k,\ell}] \in \mathbf{E}_{k^2}$. Recalling that

$$\mathbb{E}[f_n^\varphi : f_n^\varphi > \ell] + \mathbb{E}[f_n^\varphi : f_n^\varphi \leqslant \ell] = 0,$$

we have

$$\begin{aligned}\mathbb{E}[f_n^\varphi\Phi_{k,n,\ell}] &= k\,\mathbb{E}[f_n^\varphi : f_n^\varphi > \ell] + \frac{1}{k}\mathbb{E}[f_n^\varphi : f_n^\varphi \leqslant \ell] \\ &= \left(k - \frac{1}{k}\right)\mathbb{E}[f_n^\varphi : f_n^\varphi > \ell],\end{aligned}$$

and by (4.12)

$$\begin{aligned}\limsup_{n\to\infty}\mathbb{E}[f_n^\varphi\Phi_{k,n,\ell}] &\leqslant C(\theta,\varphi)\limsup_{n\to\infty}\left[k\,\mathbb{P}[f_n^\varphi > \ell] + \frac{1}{k}\mathbb{P}[f_n^\varphi \leqslant \ell]\right] \\ &\leqslant C(\theta,\varphi)\limsup_{n\to\infty}\left[\frac{k}{\ell}\mathbb{E}[|f_n^\varphi|] + \frac{1}{k}\right].\end{aligned}$$

From this and (4.10) we obtain

$$\limsup_{n\to\infty} \mathbb{E}[f_n^\varphi : f_n^\varphi > \ell] \leqslant C(\theta,\varphi)\frac{k/\ell + 1/k}{k - 1/k}.$$

Since k is arbitrary it is clear that $\{f_n^{\varphi,+}\}$ is uniformly integrable.

Step 5. Now we have to deal with the negative parts of $\{f_n^\varphi\}$. The difficulty is that they may not be uniformly integrable, and we have to find a way to split $f_n^{\varphi,-} = -(f_n^\varphi \wedge 0)$ (possibly just along a subsequence) into uniformly integrable and "bad" parts, so that the bad part plays no role in the limit. This is done using the following lemmas.

LEMMA 4.3 *Let $\{h_n\}$, $n \geqslant 1$, be a sequence of nonnegative functions and*

$$\sup_n \mathbb{E}[h_n] \leqslant C.$$

Then there is a subsequence $\{n_j\}$, $j \geqslant 1$, such that $h_{n_j} = \widehat{h}_{n_j} + r(h_{n_j})$, where

(i) $\widehat{h}_{n_j} = h_{n_j} \mathbb{1}_{\{h_{n_j} \leqslant \ell_j\}}$, *$\ell_j \to \infty$ as $j \to \infty$, and $\{\widehat{h}_{n_j}\}$, $j \geqslant 1$, is uniformly integrable;*

(ii) *remainder terms $r(h_{n_j})$ converge to* 0 *in probability as $j \to \infty$.*

Properties (i) *and* (ii) *continue to hold if we replace ℓ_j by any slower-growing sequence $\ell'_j \leqslant \ell_j$ such that $\ell'_j \to \infty$ as $j \to \infty$.*

LEMMA 4.4 *Let $\{h_n\}$, $n \geqslant 1$, be a sequence of nonnegative functions such that*

$$\sup_n \mathbb{E}[h_n] \leqslant C,$$

and let $\psi \in C_0^\infty(\mathbb{R}^d)$ be a nonnegative mollifier. Apply Lemma 4.3 *to $\{h_n^\psi\}$. Then the convergence of the remainder terms $r(h_{n_j}^\psi)$ to* 0 *is locally uniform in x, i.e., for every $R > 0$ and $\varepsilon > 0$*

$$\mathbb{P}\Big\{\omega : \sup_{|x|\leqslant R} r(h_{n_j}^\psi)(\tau_{(0,x)}\omega) \geqslant \varepsilon\Big\} \to 0 \quad \text{as } j \to \infty.$$

The proofs of Lemma 4.3 and Lemma 4.4 are elementary and are provided in the appendix.

Return now to f_n^φ. Let $\varphi = \varphi_1 * \varphi_2$, where φ_1 satisfies (4.8). By Lemma 4.3 we can write (considering every subsequence as a whole sequence)

$$f_n^{\varphi_1} = m_n - b_n,$$

where $m_n = f_n^{\varphi_1,+} - \widehat{f_n^{\varphi_1,-}}$ are uniformly integrable, $b_n = r(f_n^{\varphi_1,-})$, and $b_n \to 0$ in probability. Taking a convolution with φ_2 in this decomposition, we get

$$f_n^\varphi = m_n^{\varphi_2} - b_n^{\varphi_2}.$$

It is not hard to show that $m_n^{\varphi_2}$, $n \geqslant 1$, are still uniformly integrable. Apply Lemma 4.4 to $b_n^{\varphi_2}$ and get

$$b_n^{\varphi_2} = \widehat{b_n^{\varphi_2}} + r(b_n^{\varphi_2}),$$

where $\widehat{b_n^{\varphi_2}}$ are uniformly integrable and $r(b_n^{\varphi_2}) \to 0$ in probability locally uniformly with respect to the spatial shifts, i.e.,

$$\sup_{|x|\leqslant R} r(b_n^{\varphi_2})(\tau_{(0,x)}\omega) \to 0$$

in probability as $n \to \infty$. Putting the two decompositions together, we obtain

$$f_n^{\varphi} = k_n - r_n,$$

where $k_n = m_n^{\varphi_2} - \widehat{b_n^{\varphi_2}}$ and $r_n = r(b_n^{\varphi_2})$. This decomposition is "stable" in the following sense: further mollification and decomposition of the remainder terms r_n will not contribute any additional uniformly integrable piece.

Step 6. Let $g \in L^{\alpha}(\Omega; \mathbb{R}^d)$ be a weak limit point of $\{g_n\}$, $n \geqslant 1$. Fix a small $r > 0$. Take $\varphi = \varphi_1 * \varphi_2$, where φ_i, $i = 1, 2$, are mollifiers, $\varphi \equiv 0$ outside of $B(0, r)$, and φ_1 satisfies (4.8). According to Step 4, the positive part of f_n^{φ} is uniformly integrable. Decompose the negative part of f_n^{φ} as in Step 5 and let $\tilde{f}^{\varphi} \in L^1(\Omega, \mathbb{P})$ be a weak limit point of k_n. Then $\mathbb{E}[\tilde{f}^{\varphi}] \geqslant 0$. It is convenient to write $\tilde{f}^{\varphi}$ as $f_{\varphi} + c$ where $c \geqslant 0$ is a constant and $\mathbb{E}[f_{\varphi}] = 0$. We can drop c and inequality (4.4) will still hold. This gives (iii). It is clear also that f_{φ} and g satisfy (i) and (ii).

To show part (iv) we use (iii) and the lower bound on H to get

$$f_{\varphi}(\omega) + \frac{1}{2} \operatorname{div} g^{\varphi}(\omega) + c_1 \int |g(\tau_{(0,x)}\omega)|^{\alpha} \varphi(x)\,dx \leqslant \overline{H}(\theta) + c_2. \tag{4.13}$$

Applying the inequality

$$|\langle p, q\rangle| \leqslant \varepsilon |p|^{\alpha} + C_{\varepsilon} |q|^{\alpha'}$$

with $p = g(\tau_{(0,x)}\omega)$, $q = (\nabla\varphi_1)/\varphi_1$, and $\varepsilon = c_1$ and recalling that φ_1 satisfies (4.8), we obtain

$$\begin{aligned}
&\frac{1}{2} \operatorname{div} g^{\varphi_1}(\omega) + c_1 \int |g(\tau_{(0,x)}\omega)|^{\alpha} \varphi_1(x)\,dx \\
&\quad = c_1 \int |g(\tau_{(0,x)}\omega)|^{\alpha} \varphi_1(x)\,dx - \frac{1}{2} \int \langle g(\tau_{(0,x)}\omega), \nabla\varphi_1(x)\rangle\,dx \\
&\quad \geqslant \frac{1}{2}\left(c_1 \int |g(\tau_{(0,x)}\omega)|^{\alpha} \varphi_1(x)\,dx - C_{\varepsilon}[C(\varphi_1)]^{\alpha'}\right).
\end{aligned}$$

Convoluting every term of the above inequality with φ_2, substituting in (4.13), and applying Hölder's inequality, we get (iv). $\square$

COROLLARY 4.5 *Let H satisfy* (H1)–(H3). *For each $\theta \in \mathbb{R}^d$ there is a positive function $\nu(\delta) \to 0$ as $\delta \to 0$ and a family of functions $(f_\delta, g_\delta) : \Omega \to \mathbb{R} \times \mathbb{R}^d$, $f_\delta \in L^1(\Omega, \mathbb{P})$, $g_\delta \in L^\alpha(\Omega, \mathbb{P})$, $\delta \in (0, \delta_0)$, such that*

(i) $\mathbb{E}(f_\delta, g_\delta) = (0,0)$, $\nabla \times g_\delta = 0$, *and* $(D_t, \nabla) \times (f_\delta, g_\delta) = 0$;

(ii) $f_\delta(\tau_{(\cdot,\cdot)}\omega), g_\delta(\tau_{(\cdot,\cdot)}\omega) \in C^\infty(\mathbb{R}^{d+1}; \mathbb{R}^{d+1})$ *for $\mathbb{P}$-a.e. ω;*

(iii) *there is a constant $K_\delta > 0$, which depends only on δ, c_1, 4α, and $\overline{H}(\theta)$, such that for almost every ω with respect to $\mathbb{P}$*

$$f_\delta(\tau_{(t,x)}\omega) + \frac{c_1}{2}|g_\delta(\tau_{(t,x)}\omega)|^\alpha \leqslant K_\delta \quad \textit{for all } (t,x) \in \mathbb{R}^{d+1};$$

(iv) *for almost every ω with respect to $\mathbb{P}$*

$$f_\delta(\tau_{(t,x)}\omega) + \operatorname{div} g_\delta(\tau_{(t,x)}\omega) + H(\theta + g_\delta(\tau_{(t,x)}\omega), \tau_{(t,x)}\omega) \leqslant \overline{H}(\theta) + \nu(\delta)$$

on $\mathbb{R}^{d+1}$.

PROOF: Assumption (H3) allows us to replace

$$\int H(\theta + g(\tau_{(0,x)}\omega), \tau_{(0,x)}\omega)\varphi(x)dx$$

in (4.4) with $H(\theta + g^\varphi(\omega), \omega)$ modulo a small error. See lemma 6.3 of [14] for details. To gain the regularity with respect to space-time shifts, we can mollify one more time in both t and x and again bring the mollification inside H. □

Now we are ready to construct approximate supersolutions V_δ. Set

$$V_\delta(t,x,\omega) = \int_{(0,0)\to(t,x)} f_\delta(\tau_{(s,x(s))}\omega)ds + \langle g_\delta(\tau_{(s,x(s))}\omega), dx(s)\rangle, \tag{4.14}$$

where $x(s)$ is any smooth path such that $x(0) = 0$ and $x(t) = x$. Since $(D_t, \nabla) \times (f_\delta, g_\delta) = 0$, the integral does not depend on the path. Main properties of V_δ are summarized in the next lemma.

LEMMA 4.6 *Let $V_\delta(t,x,\omega)$ be given by* (4.14). *Then*

(i) *for almost every ω with respect to $\mathbb{P}$ and all $t, s \in \mathbb{R}$, $x, y \in \mathbb{R}^d$,*

$$V_\delta(0,0,\omega) = 0, \qquad V_\delta(s,y,\tau_{(t,x)}\omega) = V_\delta(t+s, x+y, \omega) - V_\delta(t,x,\omega);$$

(ii) $\partial_t V_\delta(t,x,\omega) = f_\delta(\tau_{(t,x)}\omega)$, $\nabla V_\delta(t,x,\omega) = g_\delta(\tau_{(t,x)}\omega)$, *and $V_\delta(t,x,\omega)$ satisfies* (4.1);

(iii) *for all $t, s \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$*

$$\|V_\delta(t,x,\cdot) - V_\delta(s,y,\cdot)\|_1 \leqslant C(\delta)(|t-s| + |x-y|). \tag{4.15}$$

(iv) *There are positive constants* $C_1 = C_1(d)$ *and* $C_2 = C_2(\delta, c_1, \alpha, \overline{H}(\theta))$ *such that with probability* 1 *for all* $(t, x) \in \mathbb{R}^{d+1}$ *and* $\eta > 0$,

$$V_\delta(t, x, \omega) \leqslant C_1 \eta^{-(d+1)} \int_0^\eta \int_{|y| \leqslant s} V_\delta(t - s, x - y, \omega) dy\, ds + C_2 \eta,$$

$$V_\delta(t, x, \omega) \geqslant C_1 \eta^{-(d+1)} \int_0^\eta \int_{|y| \leqslant s} V_\delta(t + s, x + y, \omega) dy\, ds - C_2 \eta.$$

PROOF: Parts (i)–(iii) are immediate from (4.14) and Corollary 4.5. We only need to prove (iv).

Let δ, t, x, ω be fixed and (s, y) be an arbitrary vector in $\mathbb{R} \times \mathbb{R}^d$. Consider

$$W(s, y) := V_\delta(t + s, x + sy, \omega).$$

Then by part (iii) of Corollary 4.5

$$\partial_s W(s, y) = f_\delta + \langle g_\delta, y \rangle \leqslant K_\delta - \frac{c_1}{2} |g_\delta|^\alpha + \langle g_\delta, y \rangle \leqslant K_1 + K_2 |y|^{\alpha'}.$$

Integrating we get for $s \geqslant 0$

$$V_\delta(t + s, x + sy, \omega) - V_\delta(t, x, \omega) \leqslant [K_1 + K_2 |y|^{\alpha'}] s. \tag{4.16}$$

In particular, for $s > 0$

$$V_\delta(t + s, x + y, \omega) - V_\delta(t, x, \omega) \leqslant \left[K_1 + K_2 \left(\frac{|y|}{s} \right)^{\alpha'} \right] s.$$

Notice that (4.16) leads to the following a.s. estimates for fixed (t, x) and $\eta > 0$:

$$\begin{aligned} V_\delta(t, x, \omega) &\leqslant \min_{\substack{0 \leqslant s \leqslant \eta \\ |y| \leqslant 1}} V_\delta(t - s, x - sy, \omega) + (K_1 + K_2)\eta \\ &\leqslant C_1 \eta^{-(d+1)} \int_0^\eta \int_{|y| \leqslant s} V_{\delta,\varepsilon}(t - s, x - y, \omega) dy\, ds + C_2 \eta. \end{aligned}$$

The other inequality is obtained in a similar way. □

5 Upper Bound

We start with a lemma (see [14, lemma 4.1] for a proof). Let $c \in \mathcal{C}$ and $y_\varepsilon(s)$, $s \in [t, T]$, be a sample path of a diffusion with generator $\mathbf{A}_s^\varepsilon$ (see (2.13)), and

$$\xi_\varepsilon(t) = \int_t^T L(c(s, y_\varepsilon(s)), \tau_{(s/\varepsilon, y_\varepsilon(s)/\varepsilon)} \omega) ds. \tag{5.1}$$

LEMMA 5.1 *Assume* (L1) *and* (2.2). *Then in the variational formula* (2.14), *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*: *there is a* $C > 0$ *that depends only on the constants in* (L1) *and* (2.2) *such that*

$$\sup_{x,\omega} E^{Q^{c,\varepsilon}_{t,x}} |\xi_\varepsilon(t)| \leqslant C\big(T - t + \sqrt{\varepsilon(T-t)}\big). \tag{5.2}$$

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

$$\sup_{x,\omega} E^{Q^{c,\varepsilon}_{t,x}}\left[\int_t^T |c(s, y_\varepsilon(s))|^{\beta'} ds\right] \leqslant C\big(T - t + \sqrt{\varepsilon(T-t)}\big). \tag{5.3}$$

The following theorem is the main result of this section:

THEOREM 5.2 *Assume that* $H(p,\omega)$ *satisfies* (H1)–(H3) *and that the terminal condition* $U(x)$ *satisfies* (2.2). *Let* $u(t,x)$ *be the unique solution of* (1.2)–(1.3) *with* $\overline{H}$ *given by* (2.8). *Then with probability* 1 *for every* $\ell > 0$

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

PROOF: By Lemma 5.1 the supremum in (2.14) can be restricted to controls $c \in \mathcal{C}^*$. Estimate (5.3) implies that $y_\varepsilon(T)$ is uniformly integrable with respect to $\{Q^{\varepsilon,c}_{t,x} : |x| \leqslant \ell,\ 0 \leqslant t \leqslant T,\ 0 < \varepsilon \leqslant 1,\ c \in \mathcal{C}^*\}$. Combining this with (5.2) we get

$$\sup_{|x|\leqslant\ell}\ \sup_{0\leqslant t\leqslant T}\ \sup_{c\in\mathcal{C}^*}\ \sup_{0<\varepsilon\leqslant 1} E^{Q^{\varepsilon,c}_x}(|y_\varepsilon(t)| + |\xi_\varepsilon(t)|) < \infty. \tag{5.4}$$

Step 1. Without loss of generality we can assume that $(T-t) \geqslant r$ (see Lemma 3.1). Then using the uniform integrability we get that for every $\delta > 0$ there are $M = M(\delta)$ and $A = A(\delta, r)$ such that

$$\begin{aligned}
&u_\varepsilon(t,x,\omega) - u(t,x)\\
&\quad= \sup_{c\in\mathcal{C}^*} E^{Q^{c,\varepsilon}_{t,x}}(f(y_\varepsilon(T)) - \xi_\varepsilon(t)) - \sup_{y\in\mathbb{R}^d}\left[f(y) - (T-t)\overline{L}\left(\frac{y-x}{T-t}\right)\right]\\
&\quad\leqslant \sup_{c\in\mathcal{C}^*} E^{Q^{c,\varepsilon}_{t,x}}\left[(f(y_\varepsilon(T)) - \xi_\varepsilon(t))\mathbb{1}_{|y_\varepsilon(T)-x|\leqslant M}\right] + \delta\\
&\qquad - \sup_{y\in\mathbb{R}^d}\left[f(y) - (T-t)\overline{L}\left(\frac{y-x}{T-t}\right)\right]\\
&\quad\leqslant \sup_{c\in\mathcal{C}^*} E^{Q^{c,\varepsilon}_{t,x}}\left[\left((T-t)\overline{L}\left(\frac{y_\varepsilon(T)-x}{T-t}\right) - \xi_\varepsilon(t)\right)\mathbb{1}_{|y_\varepsilon(T)-x|\leqslant M}\right] + \delta\\
&\quad= \sup_{c\in\mathcal{C}^*} E^{Q^{c,\varepsilon}_{t,x}}\Big[\sup_{|\theta|\leqslant A}\big(\langle\theta, (y_\varepsilon(T)-x)\rangle - (T-t)\overline{H}(\theta) - \xi_\varepsilon(t)\big)\mathbb{1}_{|y_\varepsilon(T)-x|\leqslant M}\Big] + \delta\\
&\quad= \sup_{c\in\mathcal{C}^*} \int\limits_{\{|y-x|\leqslant M\}\times\mathbb{R}} \sup_{|\theta|\leqslant A}\big(\langle\theta, y-x\rangle - (T-t)\overline{H}(\theta) - \xi\big)\mu_\varepsilon(dy,d\xi) + \delta,
\end{aligned}$$

where μ_ε is the distribution of $(y_\varepsilon(T), \xi_\varepsilon(t))$ on $\mathbb{R}^d \times \mathbb{R}$ under $Q^{c,\varepsilon}_{t,x}$. Since $\overline{H}$ is continuous, the supremum over θ can be restricted to the set $\{|\theta| \leqslant A, \theta \in \mathbb{Q}\}$, where $\mathbb{Q}$ is the countable set of points in $\mathbb{R}^d$ with rational coordinates. Since $L(\cdot, \omega)$ is superlinear, the integration can be further restricted to the set $\{(y, \xi) : |y - x| \leqslant M,\ \xi \leqslant B\}$ for a sufficiently large B (the proof is similar to that for lemma 4.1 in [14]). Therefore, to obtain an almost sure, locally uniform upper bound, it is enough to show that for each $\theta \in \mathbb{R}^d$, $\ell, \eta > 0$, and all large enough M and B,

$$\lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{|x|\leqslant\ell} \sup_{0\leqslant t\leqslant T} Q^{\varepsilon,c}_{t,x}\big\{\langle\theta, y_\varepsilon(T) - x\rangle - (T-t)\overline{H}(\theta) - \xi_\varepsilon(t) \geqslant \eta; A_{t,x}\big\} = 0, \tag{5.5}$$

ω-a.s. Here $A_{t,x} = \{|y_\varepsilon(T) - x| \leqslant M, \xi_\varepsilon(t) \leqslant B\}$.

Step 2. Since all computations below are valid for all $\omega \in \Omega'$, $\mathbb{P}(\Omega') = 1$, we fix $\omega \in \Omega'$ and drop it from the notation. Define

$$V_{\delta,\varepsilon}(t, x) = \varepsilon V_\delta(t/\varepsilon, x/\varepsilon). \tag{5.6}$$

By part (ii) of Lemma 4.6, $V_{\delta,\varepsilon}(t, x)$ satisfies

$$\partial_t V_{\delta,\varepsilon} + \frac{\varepsilon}{2}\Delta V_{\delta,\varepsilon} + H(\theta + \nabla V_{\delta,\varepsilon}, \tau_{(t/\varepsilon, x/\varepsilon)}\omega) \leqslant \overline{H}(\theta) + \nu(\delta), \quad (t, x) \in \mathbb{R} \times \mathbb{R}^d.$$

From Itô's formula for $V_{\delta,\varepsilon}(t, x) + \langle\theta, x\rangle$ and the above inequality we get

$$\begin{aligned} &\langle\theta, y_\varepsilon(T) - x\rangle - (T-t)\overline{H}(\theta) - \xi_\varepsilon(t) \\ &\quad \leqslant V_{\delta,\varepsilon}(t, x) - V_{\delta,\varepsilon}(T, y_\varepsilon(T)) \\ &\qquad + \sqrt{\varepsilon}\int_t^T \langle\nabla V_{\delta,\varepsilon}(s, y_\varepsilon(s)) + \theta, dB(s)\rangle + (T-t)\nu(\delta). \end{aligned} \tag{5.7}$$

We also used the fact that $\langle c, p\rangle \leqslant H(p, \omega) + L(c, \omega)$. Therefore, it is enough to show that for each $\theta \in \mathbb{R}^d$, $\ell, \eta > 0$, all sufficiently large M and B, and small $\delta > 0$,

$$\begin{aligned} \lim_{\varepsilon\to 0} \sup_{c\in\mathcal{C}^*} \sup_{|x|\leqslant\ell} \sup_{0\leqslant t\leqslant T} Q^{\varepsilon,c}_{t,x}\bigg\{&\sqrt{\varepsilon}\int_t^T \langle\nabla V_{\delta,\varepsilon}(s, y_\varepsilon(s)), dB(s)\rangle \\ &+ V_{\delta,\varepsilon}(t, x) - V_{\delta,\varepsilon}(T, y_\varepsilon(T)) \geqslant \eta; A_{t,x}\bigg\} = 0. \end{aligned} \tag{5.8}$$

Observe that again by Itô's formula

$$
\begin{aligned}
V_{\delta,\varepsilon}(T, y_\varepsilon(T)) - V_{\delta,\varepsilon}(t,x) - \sqrt{\varepsilon}\int_t^T \langle \nabla V_{\delta,\varepsilon}(s, y_\varepsilon(s)), dB(s)\rangle \\
= \int_t^T \partial_s V_{\delta,\varepsilon}(s, y_\varepsilon(s)) + \langle \nabla V_{\delta,\varepsilon}(s, y_\varepsilon(s)), c(s, y_\varepsilon(s))\rangle \\
+ \frac{\varepsilon}{2}\Delta V_{\delta,\varepsilon}(s, y_\varepsilon(s))\, ds.
\end{aligned} \tag{5.9}
$$

Setting

$$
F_{\delta,\varepsilon}(s,y) = \partial_s V_{\delta,\varepsilon}(s,y) + \langle \nabla V_{\delta,\varepsilon}(s,y), c(s,y)\rangle + \frac{\varepsilon}{2}\Delta V_{\delta,\varepsilon}(s,y),
$$

we conclude from (5.8) and (5.9) that the relation

$$
\lim_{\varepsilon\to 0}\sup_{c\in\mathcal{C}^*}\sup_{|x|\leqslant \ell}\sup_{0\leqslant t\leqslant T} E^{Q^{\varepsilon,c}_{t,x}}\left(\left[\int_t^T F_{\delta,\varepsilon}(s, y_\varepsilon(s))ds\right]^2; A_{t,x}\right) = 0
$$

will imply the desired upper bound. The proof follows step 2 of the proof assuming H4 in [14, theorem 2.2].

To complete the argument, we only need to show that for every $\delta, a > 0$

$$
\lim_{\varepsilon\to 0}\sup_{\substack{0\leqslant t\leqslant T\\ |x|\leqslant a}} |V_{\delta,\varepsilon}(t,x,\omega)| = 0 \quad \mathbb{P}\text{-a.s.} \tag{5.10}
$$

Choosing $\eta = \delta/\varepsilon$ and rescaling the integrals in part (iv) of Lemma 4.6, we get

$$
|V_{\delta,\varepsilon}(t,x,\omega)| \leqslant C_1\delta^{-(d+1)}\int_{|s|\leqslant\delta}\int_{|y|\leqslant|s|} |V_{\delta,\varepsilon}(t+s, x+y, \omega)|dy\, ds + C_2\delta.
$$

Therefore,

$$
\sup_{\substack{0\leqslant t\leqslant T\\ |x|\leqslant a}} |V_{\delta,\varepsilon}(t,x,\omega)| \leqslant C_1\delta^{-(d+1)}\int_{|t|\leqslant T+\delta}\int_{|x|\leqslant a+\delta} |V_{\delta,\varepsilon}(t,x,\omega)|dx\, dt + C_2\delta,
$$

and (5.10) is an immediate consequence of the scaling relation (5.6) and the following ergodic theorem. Note that, in this context, the variable t does not play a special role, and we only need to apply the ergodic theorem in $\mathbb{R}^{d+1}$. □

THEOREM 5.3 (Ergodic Theorem) *Assume that* $\{\tau_x : x \in \mathbb{R}^d\}$ *is an ergodic-measure-preserving action on* $(\Omega, \Sigma, \mathbb{P})$. *Let* $f : \Omega \to \mathbb{R}^d$ *satisfy the following conditions*:

$$
f \in L^1(\Omega, \mathbb{P}), \quad \mathbb{E}[f] = 0, \quad \textit{and} \quad \nabla\times f = 0,
$$

in the sense that for smooth test functions φ

$$
\int [f_i(\tau_x\omega)\partial_j\varphi(x) - f_j(\tau_x\omega)\partial_i\varphi(x)]dx = 0 \quad \textit{a.e. } \mathbb{P}; \tag{5.11}
$$

i.e., f is formally a gradient. Define the normalized integral $F : \mathbb{R}^d \to L^1(\Omega, \mathbb{P})$ by

$$F(x,\omega) = \int_0^t \langle f(\tau_{x(s)}\omega), dx(s)\rangle,$$

where $x(s)$ is a smooth path in $\mathbb{R}^d$ such that $x(0) = 0$ and $x(t) = x$. Then for every bounded set $D \subset \mathbb{R}^d$

$$\lim_{\varepsilon\to 0} \varepsilon \int_D |F(x/\varepsilon,\omega)|dx = 0 \quad \textit{a.e. } \mathbb{P}.$$

This theorem, which can be viewed as one possible generalization of the usual Birkhoff ergodic theorem, does not seem to exist in the literature. We will outline a standard proof in the next section that reduces it to Wiener's multidimensional version of Birkhoff's theorem through a series of steps.

6 Proof of Theorem 5.3

Wiener's multidimensional ergodic theorem states that if $f \in L^1(\Omega, \mathbb{P})$ and $\mathbb{E}[f] = 0$, then

$$\lim_{\varepsilon\to 0} \int_D f(\tau_{x/\varepsilon}\omega)dx = 0 \quad \text{a.e. } \mathbb{P}$$

for a large class of sets that includes all bounded boxes with sides parallel to the axes. It is not hard to strengthen it and show that for any continuous function g on $\mathbb{R}^d$

$$\lim_{\varepsilon\to 0} \int_D f(\tau_{x/\varepsilon}\omega)g(x)dx = 0 \quad \text{a.e. } \mathbb{P}.$$

LEMMA 6.1 *If f is as in Theorem* 5.3*, then*

$$\lim_{\varepsilon\to 0} \varepsilon \int_D F(x/\varepsilon,\omega)dx = 0 \quad \textit{a.e. } \mathbb{P},$$

where D is the unit cube $[0,1]^d$.

PROOF: We will prove by induction on k that

$$\lim_{\varepsilon\to 0} \varepsilon \int_{D_k} F(x/\varepsilon,\omega)dx_1\cdots dx_k = 0 \quad \text{a.e. } \mathbb{P}$$

where $D_k = \{x : x \in D, x_j = 0 \text{ for } j \geqslant k+1\}$. For $k = 1$, this means proving the almost sure convergence

$$\lim_{\varepsilon\to 0} \varepsilon \sup_{0\leqslant a\leqslant 1} \left| \int_0^{a\varepsilon^{-1}} f_1(\tau_{ye_1}\omega)dy \right| = 0 \quad \text{a.e. } \mathbb{P}.$$

The ergodic theorem will almost yield this except that

$$\lim_{\ell\to\infty} \frac{1}{\ell}\int_0^\ell f_1(\tau_{ye_1}\omega)dy = c(\omega)$$

may not be 0. The action of translations in just one coordinate direction may not be ergodic. The limit $c(\omega)$ may not be invariant under translations in other coordinate directions. If it is, then it is a constant and hence 0. Since $\{f_i\}$ is a gradient, an elementary calculation shows that

$$\int_0^\ell [f_1(\tau_{ye_1}\tau_x\omega) - f_1(\tau_{ye_1}\omega)]dy = \int_{0\to x} \langle f(\tau_{z(t)+\ell e_1}\omega) - f(\tau_{z(t)}\omega), dz(t)\rangle,$$

which goes to 0 if divided by ℓ, proving thereby that $c(\tau_x\omega) = c(\omega)$ for all x. The same argument shows that for $1 \leqslant i \leqslant k \leqslant d$,

$$\lim_{\varepsilon\to 0}\int_{D_k} f_i(\tau_{(x_1e_1+\cdots+x_ke_k)/\varepsilon}\,\omega)dx_1\cdots dx_k = 0,$$

and with a little more work (Riemann sum approximation), for any choice of $1 \leqslant i \leqslant k \leqslant d$, continuous g on $\mathbb{R}^k$, and box $D_k \subset \mathbb{R}^k$,

$$\lim_{\varepsilon\to 0}\int_{D_k} f_i(\tau_{(x_1e_1+\cdots+x_ke_k)/\varepsilon}\,\omega)g(x_1,\ldots,x_k)dx_1\cdots dx_k = 0.$$

Now we do the induction on k. Assume that for any box $D_k \subset \mathbb{R}^k$,

$$\lim_{\varepsilon\to 0}\varepsilon\int_{D_k} F(x_1/\varepsilon,\ldots,x_k/\varepsilon,0,\ldots,0,\omega)dx_1\cdots dx_k = 0.$$

We can integrate by parts and rewrite

$$\varepsilon\int_{D_{k+1}} F(x_1/\varepsilon,\ldots,x_k/\varepsilon,x_{k+1}/\varepsilon,0,\ldots,0,\omega)dx_1\cdots dx_k\,dx_{k+1}$$

as

$$\int_{D_{k+1}} f_{k+1}(\tau_{(x_1e_1+\cdots+x_{k+1}e_{k+1})/\varepsilon}\,\omega)(1-x_{k+1})dx_1\cdots dx_k\,dx_{k+1} + \varepsilon\int_{D_k} F(x_1/\varepsilon,\ldots,x_k/\varepsilon,0,\ldots,0)dx_1\cdots dx_k.$$

This completes the induction. □

LEMMA 6.2 *Assume that* $f \in L^\infty(\Omega,\mathbb{P})$. *Then*

$$\lim_{\varepsilon\to 0}\varepsilon\int_D |F(x/\varepsilon,\omega)|dx = 0.$$

PROOF: Clearly,

$$\varepsilon|F(x/\varepsilon,\omega) - F(y/\varepsilon,\omega)| \leqslant C|x-y|$$

and $F(0,\omega) \equiv 0$. For fixed ω, if $F^*(x)$ is any weak limit point of $\varepsilon F(x/\varepsilon,\omega)$, it is enough to show that $F^*(x)$ is a constant. Therefore it suffices to show that, for every smooth $\psi : \mathbb{R}^d \to \mathbb{R}^d$ with compact support,

$$\lim_{\varepsilon\to 0} \int \varepsilon F(x/\varepsilon,\omega)\,\mathrm{div}\,\psi(x)dx = 0$$

a.e. $\mathbb{P}$. We can integrate by parts and reduce it to

$$\lim_{\varepsilon\to 0} \int \langle f(\tau_{x/\varepsilon}\omega), \psi(x)\rangle dx = 0,$$

which is an easy consequence of Wiener's multidimensional ergodic theorem. □

LEMMA 6.3 *Let* $f \in L^1(\Omega,\mathbb{P})$ *with* $\mathbb{E}[f] = 0$ *satisfying* (5.11). *Then*

$$\limsup_{\varepsilon\to 0} \varepsilon \int_D |F(x/\varepsilon,\omega)|dx \leqslant \frac{1}{2}\,\mathbb{E}\Big[\sum_{i=1}^{d} |f_i|\Big] \quad a.e.\ \mathbb{P}. \tag{6.1}$$

PROOF: Let $D = [0,1]^d$, and F be an integrable function on D whose weak gradient $\nabla F = f \in L^1(D;\mathbb{R}^d)$. Then

$$\begin{aligned}
&\int_{D\times D} |F(x) - F(y)|dx\,dy \\
&\quad\leqslant \sum_{i=1}^{d} \int_{D\times D} |F(\ldots, y_{i-1}, x_i, x_{i+1},\ldots) - F(\ldots, y_{i-1}, y_i, x_{i+1},\ldots)|dx\,dy \\
&\quad= \sum_{i=1}^{d} \int_{D\times D} \left|\int_{y_i}^{x_i} f_i(\ldots, y_{i-1}, s, x_{i+1},\ldots)ds\right| dx\,dy \\
&\quad\leqslant 2\sum_{i=1}^{d} \int_{D} s(1-s)|f_i(\ldots, y_{i-1}, s, x_{i+1},\ldots)|dy_1\cdots dy_{i-1}\,ds\,dx_{i+1}\cdots dx_d \\
&\quad\leqslant \frac{1}{2}\int_D \sum_{i=1}^{d} |f_i(x)|dx.
\end{aligned}$$

For a.e. ω

$$\begin{aligned}
&\int_D \varepsilon|F(x/\varepsilon,\omega)|dx \\
&\quad\leqslant \varepsilon\int_D \left|F(x/\varepsilon,\omega)-\int_D F(y/\varepsilon,\omega)dy\right|dx+\varepsilon\left|\int_D F(y/\varepsilon,\omega)dy\right| \\
&\quad\leqslant \varepsilon\int_{D\times D} |F(x/\varepsilon,\omega)-F(y/\varepsilon,\omega)|dx\,dy+\varepsilon\left|\int_D F(y/\varepsilon,\omega)dy\right| \\
&\quad\leqslant \frac{1}{2}\sum_{i=1}^{d}\int_D |f_i(x/\varepsilon,\omega)|dx+\varepsilon\left|\int_D F(y/\varepsilon,\omega)dy\right|.
\end{aligned}$$

Applying Wiener's ergodic theorem and Lemma 6.1, we get the result. □

Now we return to the proof of Theorem 5.3.

PROOF: In view of Lemmas 6.1 and 6.2, we only need to approximate gradients $f \in L^1(\Omega,\mathbb{P})$ by those in $L^\infty(\Omega,\mathbb{P})$. The approximation itself has to be done with some care. We begin with the relation

$$\frac{\partial}{\partial\lambda}[\lambda^d\varphi(\lambda x)] = d\lambda^{d-1}\varphi(\lambda x)+\lambda^d\sum_j x_j\left(\frac{\partial\varphi}{\partial x_j}\right)(\lambda x) = \frac{1}{\lambda}\sum_{j=1}^{d}\frac{\partial\psi_j^\lambda}{\partial x_j}(x),$$

where $\psi_j^\lambda(x) = x_j\lambda^d\varphi(\lambda x)$ and φ is a mollifier with compact support. This implies that for any $\lambda_1,\lambda_2>0$ the function $\lambda_1^d\varphi(\lambda_1 x)-\lambda_2^d\varphi(\lambda_2 x)$ is the divergence of a nice object. Take $\lambda_1 = n$ and $\lambda_2 = n^{-1}$ and define

$$g_n(x) = n^d\varphi(nx)-n^{-d}\varphi(n^{-1}x).$$

Then g_n can be written as $\operatorname{div}\psi^{(n)}$ for some smooth function $\psi^{(n)}:\mathbb{R}^d\to\mathbb{R}^d$. Let $f_j^{(n)}(\omega) = (f_j * g_n)(\omega)$. Using (5.11) it is easy to check that $f^{(n)} = \{f_j^{(n)}\}$, $j = 1,\dots,d$, is a gradient, namely,

$$f^{(n)}(\omega) = \nabla h_n(\omega)\quad\text{where } h_n(\omega) = -\sum_{j=1}^{d}(f_j*\psi_j^{(n)})(\omega)\in L^1(\Omega,\mathbb{P}).$$

Moreover, as $n\to\infty$ each $f_j^{(n)}$ converges in $L^1(\Omega,\mathbb{P})$ to f_j because, by the ergodic theorem, the large-scale average corresponding to λ_2 tends to 0. Now it is routine to truncate h_n, mollify the truncation, and take gradients to get nice bounded gradients that approximate f in $L^1(\Omega,\mathbb{P})$. □

Appendix: Proofs of Technical Lemmas

PROOF OF LEMMA 4.3: For $\ell = 1, 2, \ldots$, define $h_n^\ell = h_n \mathbb{1}_{\{h_n \leqslant \ell\}}$. Then $0 \leqslant h_n^\ell \leqslant \ell$, and we can choose a subsequence $\{n_j\}$ such that, for each ℓ, functions $h_{n_j}^\ell \in L^\infty(\Omega)$ converge weakly as $j \to \infty$ to some h^ℓ, $0 \leqslant h^\ell \leqslant \ell$. Notice that

$$\mathbb{E}\, h^\ell = \lim_{j\to\infty} \mathbb{E}\, h_{n_j}^\ell \leqslant \liminf_{j\to\infty} \mathbb{E}\, h_{n_j} \leqslant C.$$

Since the sequence $\{h^\ell\}$, $\ell = 1, 2, \ldots$, is nondecreasing, we can define $h(\omega) = \lim_{\ell\to\infty} h^\ell(\omega)$ a.s. Then $\lim_{\ell\to\infty} \mathbb{E}\, h^\ell = \mathbb{E}\, h$.

Set

$$a_j^\ell = \mathbb{E} h_{n_j}^\ell, \quad a^\ell = \mathbb{E} h^\ell, \quad \text{and} \quad a = \mathbb{E} h.$$

Then

$$a_j^\ell \to a^l \text{ as } j \to \infty \quad \text{and} \quad a^\ell \nearrow a \text{ as } \ell \to \infty,$$

and we can choose a subsequence $\{\ell_j\}$, $\ell_j \to \infty$ such that $a_j^{\ell_j} \to a$ as $j \to \infty$. It is clear that for any slower-growing sequence $\ell_j' \leqslant \ell_j$, $\ell_j' \to \infty$,

$$a_j^{\ell_j'} \to a \quad \text{as } j \to \infty.$$

We claim that $\widehat{h}_{n_j} = h_{n_j}^{\ell_j}$ and $r(h_{n_j}) = h_{n_j} - h_{n_j}^{\ell_j}$ satisfy conditions (i) and (ii).

Indeed, for each ℓ and all j such that $\ell_j \geqslant \ell$

$$\int_{\widehat{h}_{n_j} > \ell} \widehat{h}_{n_j}\, d\,\mathbb{P} = \int_{h_{n_j}^{\ell_j} > l} h_{n_j}^{\ell_j}\, d\,\mathbb{P}$$

$$= a_j^{\ell_j} - \int_{h_{n_j}^{\ell_j} \leqslant l} h_{n_j}^{\ell_j}\, d\,\mathbb{P} = a_j^{\ell_j} - \int h_{n_j}^\ell\, d\,\mathbb{P} = a_j^{\ell_j} - a_j^\ell.$$

Thus,

$$\lim_{\ell\to\infty} \lim_{j\to\infty} \int_{\widehat{h}_{n_j} > \ell} \widehat{h}_{n_j}\, d\,\mathbb{P} = 0.$$

Also, for every $\varepsilon \in (0, 1)$,

$$\mathbb{P}(r(h_{n_j}) > \varepsilon) = \mathbb{P}(h_{n_j} - h_{n_j}^{\ell_j} > \varepsilon) \leqslant \mathbb{P}(h_{n_j} \geqslant \ell_j) \leqslant \frac{\mathbb{E}\, h_{n_j}}{\ell_j} \leqslant \frac{C}{\ell_j} \to 0$$

as $j \to \infty$. This completes the proof of Lemma 4.3. □

The proof of Lemma 4.4 is based on the following observation:

LEMMA A.1 *Let $h \geqslant 0$ be a measurable function and $\psi \geqslant 0$ be a nice mollifier on $\mathbb{R}^d$ supported on the cube of side a around the origin. Then there is a $c > 0$, which depends only on ψ, such that, for any $l > 0$ and ω,*

$$h^\psi(\omega) > \ell \Rightarrow |\{x \in \mathbb{R}^d : |x| \leqslant a\sqrt{d},\ h^\psi(\tau_{(0,x)}\omega) \geqslant cl\}| \geqslant c.$$

PROOF OF LEMMA 4.4 ASSUMING LEMMA A.1: For notational convenience we assume that $\{n_j\}$ is the whole sequence: $n_j = j$. We have

$$h_j^\psi = (\widehat{h^\psi})_j + r(h_j^\psi),$$

where $(\widehat{h^\psi})_j = h_j^\psi \mathbb{1}_{\{h_j^\psi \leqslant l_j\}}$ are uniformly integrable, $l_j \to \infty$, and $r(h_j^\psi) \to 0$ in probability. We need to show that the last convergence is locally uniform in x.

Fix $R > 0$ and $\varepsilon > 0$ and suppose that $r(h_j^\psi)(\tau_{(0,x)}\omega) \geqslant \varepsilon$ for some x in the ball of radius R around the origin. This, in fact, means that $h_j^\psi(\tau_{(0,x)}\omega) > \ell_j$. From Lemma A.1 it follows that $h_j^\psi(\tau_{0,y}\omega) \geqslant c\ell_j$ for y on a set of measure at least c in the ball $|y - x| \leqslant a\sqrt{d}$, where a, c depend only on ψ. Hence, for a ball B of radius $R + a\sqrt{d}$ around the origin

$$\int_B h_j^\psi(\tau_{(0,x)}\omega)dx \geqslant c^2\ell_j.$$

Summarizing the above, we have

$$\begin{aligned}\mathbb{P}\Big(\omega : \sup_{|x|\leqslant R} r(h_j^\psi)(\tau_{(0,x)}\omega) \geqslant \varepsilon\Big) &\leqslant \mathbb{P}\bigg(\omega : \int_B h_j^\psi(\tau_{(0,x)}\omega)dx \geqslant c^2\ell_j\bigg) \\ &\leqslant \frac{1}{c^2\ell_j}\mathbb{E}\int_B h_j^\psi(\tau_{(0,x)}\omega)dx \\ &\leqslant \frac{|B|}{c^2\ell_j}\,\mathbb{E}\,h_j^\psi \leqslant \frac{C|B|}{c^2\ell_j} \to 0 \quad \text{as } j \to \infty.\end{aligned}$$

□

PROOF OF LEMMA A.1: We have $h^\psi(\omega) = \int h(\tau_{(0,x)}\omega)\psi(x)dx$, where ψ is a nice nonnegative mollifier, bounded by A and supported on the cube of side a around the origin. We consider the cube K_b of size $b = a/N$ such that the cube of side a can be partitioned into a finite number N^d of cubes of side b. We have

$$\psi(x) \leqslant A\sum_{j=1}^{N^d} \mathbb{1}_{K_b^j}(x)$$

with $\{K_b^j\}$ being suitable translates of K_b. If $h^\psi(\omega) \geqslant \ell$, then for some j

$$\int h(\tau_{(0,y)}\omega)\mathbb{1}_{K_b^j}(y)dy \geqslant \frac{\ell}{AN^d}.$$

Since ψ is nice and $\int \psi(x)dx = 1$, it is easy to find c and N such that for any j

$$|\{x \in \mathbb{R}^d : |x| \leqslant a\sqrt{d},\ c\mathbb{1}_{K_b^j+x}(\cdot) \leqslant \psi(\cdot)\}| \geqslant c.$$

We now have a lower bound

$$h^\psi(\tau_{(0,x)}\omega) = \int h(\tau_{(0,y)}\omega)\psi(y-x)dy \geqslant \frac{c\ell}{AN^d}$$

provided $\psi(y-x) \geqslant c\mathbb{1}_{K_b^j}(y)$ or $\psi(y) \geqslant c\mathbb{1}_{K_b^j+x}(y)$.

We have shown that there is a positive constant (call it again c) depending only on ψ such that, for any $\ell > 0$, if $h^\psi(\omega) \geqslant \ell$, then

$$|\{x \in \mathbb{R}^d : |x| \leqslant a\sqrt{d},\ h^\psi(\tau_{(0,x)}\omega) \geqslant c\ell\}| \geqslant c.$$

□

Acknowledgment. E. Kosygina was partially supported by the CUNY Research Foundation under PSC-CUNY awards #66512-00-35 and #67332-00-36, and by a Eugene M. Lang Junior Faculty Research Fellowship. S. R. S. Varadhan was supported by the National Science Foundation through grants DMS-0104343 and DMS-0604380.

Bibliography

[1] Alvarez, O.; Bardi, M. Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result. *Arch. Ration. Mech. Anal.* **170** (2003), no. 1, 17–61.

[2] Arisawa, M. Quasi-periodic homogenizations for second-order Hamilton-Jacobi-Bellman equations. *Adv. Math. Sci. Appl.* **11** (2001), no. 1, 465–480.

[3] Bakhvalov, N. S.; Panasenko, G. P. *Homogenisation: averaging processes in periodic media. Mathematical problems in the mechanics of composite materials.* Mathematics and Its Applications (Soviet Series), 36. Kluwer, Dordrecht, 1989.

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

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

[6] Buckdahn, R.; Ichihara, N. Limit theorem for controlled backward SDEs and homogenization of Hamilton-Jacobi-Bellman equations. *Appl. Math. Optim.* **51** (2005), no. 1, 1–33.

[7] Caffarelli, L. A.; Souganidis, P. E.; Wang, L. Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media. *Comm. Pure Appl. Math.* **58** (2005), no. 3, 319–361.

[8] 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.

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

[10] Evans, L. C. A survey of partial differential equations methods in weak KAM theory. *Comm. Pure Appl. Math.* **57** (2004), no. 4, 445–480.

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

[12] Ishii, H. Homogenization of the Cauchy problem for Hamilton-Jacobi equations. *Stochastic analysis, control, optimization and applications*, 305–324. Systems & Control: Foundations & Applications. Birkhäuser, Boston, 1999.

[13] Jikov, V. V.; Kozlov, S. M.; Oleĭnik, O. A. *Homogenization of differential operators and integral functionals*. Springer, Berlin, 1994.

[14] Kosygina, E.; Rezakhanlou, F.; Varadhan, S. R. S. Stochastic homogenization of Hamilton-Jacobi-Bellman equations. *Comm. Pure Appl. Math.* **59** (2006), no. 10, 1489–1521.

[15] Landim, C.; Olla, S.; Yau, H. T. Convection-diffusion equation with space-time ergodic random flow. *Probab. Theory Related Fields* **112** (1998), no. 2, 203–220.

[16] 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.

[17] Lions, P.-L.; Souganidis, P. E. Homogenization of degenerate second-order PDE in periodic and almost periodic environments and applications. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **22** (2005), no. 5, 667–677.

[18] Lions, P.-L.; Souganidis, P. E. Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media. *Comm. Partial Differential Equations* **30** (2005), no. 1-3, 335–375.

[19] Marchenko, V. A.; Khruslov, E. Ya. *Homogenization of partial differential equations*. Progress in Mathematical Physics, 46. Birkhäuser, Boston, 2006.

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

[21] Sznitman, A.-S. *Brownian motion, obstacles and random media*. Springer Monographs in Mathematics. Springer, Berlin, 1998.

ELENA KOSYGINA
Baruch College
Department of Mathematics
One Bernard Baruch Way
Box B6-230
New York, NY 10010
E-mail: elena_kosygina@baruch.cuny.edu

S. R. S. VARADHAN
Courant Institute
251 Mercer Street
New York, NY 10012
E-mail: varadhan@cims.nyu.edu

Received September 2006.