ESAIM: COCV 27 (2021) 6
https://doi.org/10.1051/cocv/2021001

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SINGULAR PERTURBATIONS AND OPTIMAL CONTROL OF STOCHASTIC SYSTEMS IN INFINITE DIMENSION: HJB EQUATIONS AND VISCOSITY SOLUTIONS

Andrzej Święch*

Abstract. We study a stochastic optimal control problem for a two scale system driven by an infinite dimensional stochastic differential equation which consists of "slow" and "fast" components. We use the theory of viscosity solutions in Hilbert spaces to show that as the speed of the fast component goes to infinity, the value function of the optimal control problem converges to the viscosity solution of a reduced effective equation. We consider a rather general case where the evolution is given by an abstract semilinear stochastic differential equation with nonlinear dependence on the controls. The results of this paper generalize to the infinite dimensional case the finite dimensional results of Alvarez and Bardi [*SIAM J. Control Optim.* **40** (2001/02) 1159–1188] and complement the results in Hilbert spaces obtained recently in Guatteri and Tessitore [To appear in: *Appl. Math. Optim.* (2019) https://doi.org/10.1007/s00245-019-09577-y].

Mathematics Subject Classification. 35R15, 35B25, 35Q93, 49L25, 49L20, 60H15, 93E20.

Received June 14, 2020. Accepted December 27, 2020.

1. Introduction

In this paper we study a stochastic optimal control problem for a two scale system given by an abstract stochastic differential equation

$$\begin{cases} \mathrm{d}X_\varepsilon(s) = [-A_1X_\varepsilon(s) + b(X_\varepsilon(s), Y_\varepsilon(s), a(s))]\,\mathrm{d}s + \sigma(X_\varepsilon(s), Y_\varepsilon(s), a(s))\mathrm{d}W_Q(s) \\ \mathrm{d}Y_\varepsilon(s) = \frac{1}{\varepsilon}[-A_2Y_\varepsilon(s) + g(X_\varepsilon(s), Y_\varepsilon(s), a(s))]\,\mathrm{d}s + \frac{1}{\sqrt{\varepsilon}}h(Y_\varepsilon(s), a(s))\mathrm{d}W_Q(s) \\ X_\varepsilon(t) = x \in H, Y_\varepsilon(t) = y \in H, t < s \le T \end{cases} \tag{1.1}$$

in a real separable infinite dimensional Hilbert space H, where $T > 0$ is a fixed constant and $0 \le t \le T$. Above, A_1, A_2 are linear, densely defined, maximal monotone operators in H, b, σ, g, h are appropriate nonlinear functions, $a(\cdot)$ is a control process, W_Q is a Q-Wiener process (see Sects. 2 and 3 for the precise definitions and assumptions) and $\varepsilon > 0$ is a small parameter which will be sent to zero. Thus system (1.1) consists of a "slow"

Keywords and phrases: Hamilton-Jacobi-Bellman equation, viscosity solution, stochastic optimal control, singular perturbation, two-scale system.

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.

* Corresponding author: swiech@math.gatech.edu

component X_ε and a "fast" component Y_ε. The cost functional we want to minimize has the form

$$J_\varepsilon(t,x;a(\cdot)) = \mathbb{E}\left[\int_t^T l(X_\varepsilon(s), Y_\varepsilon(s), a(s))\mathrm{d}s + f(X_\varepsilon(T))\right].$$

The value function for the problem is defined by

$$u_\varepsilon(t,x,y) = \inf_{a(\cdot)\in\mathcal{U}_t} J_\varepsilon(t,x;a(\cdot)),$$

where $\mathcal{U}_t$ is a set of admissible controls defined in Section 2. Using the dynamic programming approach and the theory of Hamilton-Jacobi-Bellman (HJB) equations, the value function u_ε should be an appropriately defined unique viscosity solution of the associated HJB equation

$$\begin{cases} \partial_t u_\varepsilon - \langle A_1 x, D_x u_\varepsilon\rangle - \frac{1}{\varepsilon}\langle A_2 y, D_y u_\varepsilon\rangle - F\left(x, y, D_x u_\varepsilon, \frac{1}{\varepsilon} D_y u_\varepsilon, D_x^2 u_\varepsilon, \frac{1}{\varepsilon} D_y^2 u_\varepsilon, \frac{1}{\sqrt{\varepsilon}} D_{xy}^2 u_\varepsilon\right) = 0 \\ u_\varepsilon(T,x,y) = f(x) \qquad \text{in } (0,T)\times H\times H, \end{cases} \tag{1.2}$$

where the Hamiltonian F is defined by (2.1). The goal of this manuscript is to use the theory of viscosity solutions in Hilbert spaces to study the asymptotic limit of u_ε as $\epsilon \to 0$ and characterize it as a viscosity solution of an effective HJB equation.

Singular perturbation problems in stochastic optimal control have been studied extensively in finite dimensional spaces by probabilistic and partial differential equation (PDE) methods. We refer to [1–3, 5–9, 15, 17, 18, 20, 22, 24, 25] and the references therein. There is also extensive body of work on a related subject of stochastic homogenization. Much less is known for problems in infinite dimensional spaces. In a recent paper [16] a two scale stochastic system driven by abstract stochastic differential equations in Hilbert spaces was studied using backward stochastic differential equations. Only additive noise case was considered in [16]. In [25] the author investigated linear-quadratic problems with bounded operators based on the analysis of the associated Ricatti equations. Singular perturbations of infinite-dimensional deterministic control systems were studied in [13] using differential inclusions. The reader can consult [23, 26] and references in [25] for an overview of applications of singular perturbations.

In this paper we want to show that the PDE viscosity solution approach of [1] can be adapted to the infinite dimensional setting by using the theory of viscosity solutions in Hilbert spaces (see [14]). Unfortunately the theory in Hilbert spaces is not as broad and flexible and, depending on a type of an unbounded operator present, every equation may require a slightly different approach. The general strategy we use here is very similar to that of [1], however there are significant technical challenges. The viscosity solution approach in finite dimension in [1] is based on the use of Barles-Perthame half-relaxed limits (see [10]). This method requires compactness which is lacking in infinite dimensional spaces and hence fails in general. The reader can consult [14], Example 3.43 for a counterexample in a Hilbert space. However it was showed in [19] (see also [14], Sect. 3.9) that a version of the method of half-relaxed limits works for some equations in a Hilbert space under certain conditions and with an appropriate notion of a discontinuous viscosity solution. Thus we stay within this framework which requires stronger coerciveness of the operator A_1 in the slow variable equation (condition (3.2)). Another obstacle is the lack of PDE techniques to prove regularity of viscosity solutions of general HJB equations in Hilbert spaces considered here as they are degenerate, fully nonlinear and contain terms with unbounded operators. Thus we resort to a mixture of stochastic and PDE methods. We assume that the operators A_1, A_2 satisfy the so called "strong" B-condition (see Asm. 3.1). We also impose a monotonicity condition (3.4) on A_2 and the coefficients of the fast variable equation which is similar to the monotonicity condition used in [16]. Conditions like this are well-known and used in the finite dimensional literature (at least in the deterministic case) and can be traced back to [7]. Condition (3.4) guarantees uniform exponential decay of second moments of the difference of two solutions of the fast variable equation (5.1) which then imply uniform Lipschitz continuity estimates for viscosity solutions w_δ of discounted infinite horizon problems (5.2) needed to obtain the representation formula for the effective Hamiltonian $\overline{F}$. It also allows to prove good continuous dependence estimates for solutions of (5.1) which are crucial in establishing the properties of the effective Hamiltonian. Such estimates for w_δ and properties of $\overline{F}$ were proved in [1] by PDE methods and were based on coerciveness or uniform ellipticity of the equation. Unfortunately such methods are not available here. Moreover there is no hope of obtaining any second derivative estimates for w_δ and thus we need to assume that the diffusion coefficient h in the fast variable equation is independent of x. The paper [1] dealt with the periodic fast variable case while our singular perturbation problem is in the whole Hilbert space and this makes the PDE arguments more difficult. However the monotonicity condition (3.4) implies (3.10), which gives some coerciveness of the operator A_2 (weaker than that of A_1). It allows to contain the fast variable and is crucial in the proofs of Theorems 5.1 and 7.2. It is similar to a condition assumed in [15] that considered singular perturbation problems in an unbounded setting in $\mathbb{R}^n$. We also define the half-relaxed limits in the proof of Theorem 7.2 as in [15].

The results of this paper generalize to the infinite dimensional case the results of [1] and complement the results obtained in Hilbert spaces in [16]. We can deal with the case of drift and diffusion coefficients depending nonlinearly on both variables x, y and the control variable, while additive noise and a special dependence on the control variable was required in [16]. On the other hand we need to assume the continuity of some coefficients with respect to weaker negative norms (see Asm. 3.1). We also have different assumptions on the operators A_1, A_2, including compactness of the operators B_1 and B_2 associated with them (see Sect. 3), and our diffusion coefficients must be Hilbert-Schmidt, while [16] dealt with cylindrical Wiener processes and bounded constant diffusion operators. However, as far as the continuity of the coefficients, we cover the case considered in [16], see Remark 7.3. Moreover, our results also apply to the case $\sigma = h = 0$ and thus cover singular perturbations for a large class of deterministic infinite dimensional equations. Various improvements and versions of our results seem possible, for instance allowing some growth of the coefficient functions, which are assumed to be bounded here, or considering special operators A_1, A_2. Such generalizations however would perhaps require some different techniques and we do not attempt to investigate them in this manuscript. Finally, we remark that we have introduced the setup with the same state space for the slow and fast variables for simplicity, as the manuscript is already very technical. However, the techniques of the paper can be easily adapted to the case where the evolution of the slow variable is in one Hilbert space H_1 and the evolution of the fast variable is in another Hilbert space H_2, with two different Wiener processes W_{Q_1} and W_{Q_2}. In particular the slow variable could then be finite dimensional. We leave these modifications to the readers.

The plan of the paper is the following. In Section 2 we present the notation used in the manuscript and the basic definitions. The assumptions are discussed in Section 3. In Section 4 we introduce various notions of viscosity solutions needed in the paper and prove the theorem about existence of a unique viscosity solution of equation (1.2). Section 5 is devoted to the analysis of discounted infinite horizon problems in the fast variable and the definition of the effective Hamiltonian. Properties of the effective Hamiltonian are analyzed in Section 6. Finally in Section 7 we prove the main result, Theorem 7.2, about convergence of the viscosity solutions u_ε of (1.2) to the viscosity solution of the effective equation (7.1).

We refer the reader to [14] for the overview of the theory of viscosity solutions of second order HJB equations in infinite dimensional Hilbert spaces and the dynamic programming approach to optimal control of stochastic PDE.

2. Notation and definitions

All Hilbert spaces in this manuscript are real and separable. We denote by $\langle \cdot, \cdot \rangle$ the inner product in H and by $|\cdot|$ its norm. For Hilbert spaces Z, W we denote by $\mathcal{L}(Z, W)$ the space of bounded linear operators from Z to W. If $Z = W$ we will just write $\mathcal{L}(Z)$. We denote by $S(Z)$ the space of self-adjoint operators in $\mathcal{L}(Z)$. For a (non necessarily bounded) operator T we denote by T^* its adjoint. If $T \in \mathcal{L}(Z)$, $T > 0$ means that it is strictly positive, that is $\langle Tx, x \rangle > 0$ for all $x \in Z, x \neq 0$ and $T \geq 0$ means that it is positive, that is $\langle Tx, x \rangle \geq 0$ for all $x \in Z$. We denote by $\mathcal{L}_2(Z, W)$ the space of Hilbert-Schmidt operators from Z to W and if $T \in \mathcal{L}_2(Z, W)$, its norm is denoted by $\|T\|_{\mathcal{L}_2(Z,W)}$. The space of nuclear (trace class) operators on Z is denoted by $\mathcal{L}_1(Z)$. If

$T \in \mathcal{L}_1(Z)$, the trace of T is defined by

$$\mathrm{Tr}(T) = \sum_{i=1}^{\infty} \langle Te_i, e_i \rangle,$$

where $\{e_i\}$ is any orthonormal basis of Z. We refer the reader to [14], Appendix B.3 for more about Hilbert-Schmidt and nuclear operators.

Let K be a real separable Hilbert space and $Q \in S(K), Q \geq 0$. We define $K_0 := Q^{\frac{1}{2}}(K)$. If $Q^{-\frac{1}{2}}$ is the pseudo-inverse of $Q^{\frac{1}{2}}$ (see [14], Def. B.1), K_0 is a separable Hilbert space equipped with the inner product $\langle x, y \rangle_{K_0} = \langle Q^{-\frac{1}{2}} x, Q^{-\frac{1}{2}} y \rangle_K$. We say that $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q \right)$ is a reference probability space if $(\Omega, \mathscr{F}, \mathbb{P})$ is a complete probability space, W_Q is a Q-Wiener process in K, $W_Q(t) = 0$, and $\{\mathscr{F}_s^t\}_{s\in[t,T]}$ is the augmented filtration generated by W_Q. We refer to [14], Section 1.2.4 and [12] for the notion of a Q-Wiener process, and to [14], Sections 1.3 and 2.2.1 for more on the concept of a reference probability space.

For a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a separable Hilbert space Z, we denote by $L^2(\Omega, \mathscr{F}, \mathbb{P}; Z)$ (or simply $L^2(\Omega; Z)$) the space of measurable Z-valued random variables ξ such that $\mathbb{E}|\xi|_Z^2 < \infty$. The Hilbert space Z is equipped with the Borel σ-field. If μ is a reference probability space, we denote by $M_\mu^2(t, T; Z)$ the space of Z-valued progressively measurable processes $Y(\cdot)$ on $[t, T]$ such that

$$|Y(\cdot)|_{M_\mu^2(t,T;Z)} := \left(\mathbb{E} \left[\int_t^T |Y(s)|_Z^2 \mathrm{d}s \right] \right)^{\frac{1}{2}} < \infty.$$

$M_\mu^2(t, T; Y)$ is a Hilbert space with the standard inner product.

The control space Λ is a Polish space. For every reference probability space μ on $[t, T]$, we define $\mathcal{U}_t^\mu$ to be the set of all progressively measurable processes $a(\cdot) : [t, T] \to \Lambda$. We define the set of admissible controls to be

$$\mathcal{U}_t := \bigcup_\mu \mathcal{U}_t^\mu,$$

where the union is taken over all reference probability spaces μ on $[t, T]$.

If W is a subset of a Hilbert space, we denote by $C(W)$ the space of continuous functions on X with values in $\mathbb{R}$ and by $C_b(W)$ the space of bounded functions in $C(W)$. We denote by $UC(W)$ the space of uniformly continuous functions on W with values in $\mathbb{R}$. If Z is a Hilbert space and I is an interval in $\mathbb{R}$, we denote by $UB_b^x(I \times Z)$ the space of functions $u \in C_b(I \times Z)$ such that $u(t, \cdot)$ is uniformly continuous on Z, uniformly in $t \in I$. We will write $C^{1,2}((0, T) \times H)$ to denote the space of functions $\varphi : (0, T) \times H \to \mathbb{R}$ (writing $\varphi(t, x)$) such that $\partial_t \varphi, D\varphi, D^2\varphi$ are continuous on $(0, T) \times H$, where $D\varphi$ and $D^2\varphi$ stand for the first and second Fréchet derivatives of φ with respect to the x-variable. We will also write $C^{1,2}((0, T) \times H \times H)$ (writing $\varphi(t, x, y)$) to denote the space of functions on $(0, T) \times H \times H$ such that $\partial_t \varphi, D\varphi, D^2\varphi$ are continuous on $(0, T) \times H \times H$, where $D\varphi$ and $D^2\varphi$ now stand for the first and second Fréchet derivatives of φ with respect to the full (x, y) variable. For such functions we will write $D_x\varphi, D_y\varphi, D_x^2\varphi, D_y^2\varphi, D_{xy}^2\varphi$ to indicate partial Fréchet derivatives of φ with respect to the variables x and y. We will write operators $P \in S(H \times H)$ as

$$P = \begin{pmatrix} X & Z \\ Z^* & Y \end{pmatrix},$$

where $X, Y \in S(H), Z \in \mathcal{L}(H)$. In this notation

$$D^2\varphi = \begin{pmatrix} D_x^2\varphi & D_{xy}^2\varphi \\ (D_{xy}^2\varphi)^* & D_y^2\varphi \end{pmatrix}.$$

We denote $B(x,r)$ to be the open ball in H centered at x with radius $r>0$. If a ball is in a different Hilbert space Z, we will write $B_Z(x,r)$. A modulus is a continuous, nondecreasing and concave function $\rho:[0,+\infty)\to[0,+\infty)$ such that $\rho(0)=0$.

We denote

$$C(x,y,a)=\frac{1}{2}(\sigma(x,y,a)Q^{\frac{1}{2}})(\sigma(x,y,a)Q^{\frac{1}{2}})^*,\quad E(y,a)=\frac{1}{2}(h(y,a)Q^{\frac{1}{2}})(h(y,a)Q^{\frac{1}{2}})^*,$$

$$D(x,y,a)=(\sigma(x,y,a)Q^{\frac{1}{2}})(h(y,a)Q^{\frac{1}{2}})^*,$$

The Hamiltonian $F:H\times H\times H\times H\times S(H\times H)\to\mathbb{R}$ is defined by

$$\begin{aligned} F(x,y,p,q,X,Y,Z)=\sup_{a\in\Lambda}\Big\{&-\mathrm{Tr}(C(x,y,a)X)-\mathrm{Tr}(E(y,a)Y)-\mathrm{Tr}(D(x,y,a)Z)\\ &-\langle b(x,y,a),p\rangle-\langle g(x,y,a),q\rangle-l(x,y,a)\Big\}. \end{aligned} \tag{2.1}$$

3. Assumptions

We recall that A_1 and A_2 are linear, densely defined, maximal monotone operators in H. Let $B_1,B_2\in S(H)$, $B_1>0,B_2>0$ be such that

$$A_1^*B_1,A_2^*B_2\quad\text{are bounded.} \tag{3.1}$$

For $\alpha\in\mathbb{R}$ we define norms

$$|x|_{1,\alpha}:=|B_1^{-\frac{\alpha}{2}}x|,\quad |x|_{2,\alpha}:=|B_2^{-\frac{\alpha}{2}}x|.$$

We denote $H_{1,-1}$ to be the completion of H with respect to the norm $|\cdot|_{1,-1}$ and $H_{2,-1}$ to be the completion of H with respect to the norm $|\cdot|_{2,-1}$. Both are Hilbert spaces with inner products $\langle x,y\rangle_{1,-1}=\langle B_1^{\frac{1}{2}}x,B_1^{\frac{1}{2}}y\rangle$ and $\langle x,y\rangle_{2,-1}=\langle B_2^{\frac{1}{2}}x,B_2^{\frac{1}{2}}y\rangle$. We also define for $\alpha>0$ the spaces $H_{1,\alpha}=B_1^{\frac{\alpha}{2}}(H)$ and $H_{2,\alpha}=B_2^{\frac{\alpha}{2}}(H)$, which are Hilbert spaces equipped with the inner products $\langle x,y\rangle_{1,\alpha}=\langle B_1^{-\frac{\alpha}{2}}x,B_1^{-\frac{\alpha}{2}}y\rangle$ and $\langle x,y\rangle_{2,\alpha}=\langle B_2^{-\frac{\alpha}{2}}x,B_2^{-\frac{\alpha}{2}}y\rangle$ respectively. We will assume (see Asm. 3.1(i)) that B_1 and B_2 are compact. For $N\geq 1, i=1,2$, let $P_{i,N}$ be the orthogonal projection in H onto the (finite-dimensional) space spanned by the eigenvectors of B_i corresponding to the eigenvalues which are greater than or equal to $\frac{1}{N}$ and let $Q_{i,N}=I-P_{i,N}$. We remark that, since B_1 and B_2 are compact, the upper/lower semicontinuity with respect to the negative norms $|\cdot|_{1,-1}$ and $|\cdot|_{2,-1}$ on bounded sets of H is equivalent to the weak sequential upper/lower semicontinuity (see [14], Lem. 3.6).

We now introduce the assumptions we will be using in the manuscript.

Assumption 3.1.

(i) The operators B_1,B_2 are compact, there exists $\lambda>0$ such that for every $x\in D(A_1^*)$

$$\langle A_1^*x,x\rangle\geq\lambda|x|_{1,1}^2 \tag{3.2}$$

and there exists $c\geq 0$ such that for all $x,y\in H$

$$\langle(A_1^*B_1+cB_1)x,x\rangle\geq|x|^2,\quad \langle(A_2^*B_2+cB_2)y,y\rangle\geq|y|^2. \tag{3.3}$$

(ii) There exists $\mu > 0$ such that for all $x \in H, y_1, y_2 \in D(A_2^*)$, $a \in \Lambda$,

$$\begin{aligned} &- \langle A_2^*(y_1 - y_2), y_1 - y_2\rangle + \langle g(x, y_1, a) - g(x, y_2, a), y_1 - y_2\rangle \\ &+ \frac{1}{2}\mathrm{Tr}\left((h(y_1, a) - h(y_2, a))Q^{\frac{1}{2}})((h(y_1, a) - h(y_2, a))Q^{\frac{1}{2}})^*\right) \\ &\qquad \le -\mu|y_1 - y_2|^2. \end{aligned} \tag{3.4}$$

(iii) $b, g : H \times H \times \Lambda \to H, \sigma : H \times H \times \Lambda \to \mathcal{L}_2(K_0, H), h : H \times \Lambda \to \mathcal{L}_2(K_0, H), l : H \times H \times \Lambda \to \mathbb{R}, f : H \to \mathbb{R}$ are bounded, continuous and there exist $L \ge 0$ and a modulus ω such that for all $x_1, x_2, y_1, y_2 \in H, a \in \Lambda$,

$$|b(x_1, y_1, a) - b(x_2, y_2, a)| \le L(|x_1 - x_2| + |y_1 - y_2|), \tag{3.5}$$

$$\begin{aligned} &\|\sigma(x_1, y_1, a) - \sigma(x_2, y_2, a)\|_{\mathcal{L}_2(K_0,H)} + \|h(y_1, a) - h(y_2, a)\|_{\mathcal{L}_2(K_0,H)} \\ &\qquad \le L(|x_1 - x_2|_{1,-1} + |y_1 - y_2|_{2,-1}), \end{aligned} \tag{3.6}$$

$$\begin{aligned} &|g(x_1, y_1, a) - g(x_2, y_2, a)| + |l(x_1, y_1, a) - l(x_1, y_2, a)| \\ &\qquad \le L(|x_1 - x_2|_{1,-1} + |y_1 - y_2|), \end{aligned} \tag{3.7}$$

$$|l(x_1, y_1, a) - l(x_2, y_1, a)| + |f(x_1) - f(x_2)| \le \omega(|x_1 - x_2|). \tag{3.8}$$

(iv) For every $x \in H$

$$\lim_{N\to\infty} \sup_{y\in H, a\in\Lambda} \mathrm{Tr}\,(C(x, y, a)Q_{1,N}) = 0. \tag{3.9}$$

We note that, since g and h are bounded, (3.4) implies that for all $x \in H, y \in D(A_2^*), a \in \Lambda$

$$\begin{aligned} -\langle A_2^* y, y\rangle &\le -\mu|y|^2 - \langle g(x, y, a) - g(x, 0, a), y\rangle \\ &- \frac{1}{2}\mathrm{Tr}\left((h(y, a) - h(0, a))Q^{\frac{1}{2}})((h(y, a) - h(0, a))Q^{\frac{1}{2}})^*\right) \le -\mu|y|^2 + C|y| + C \end{aligned}$$

which implies

$$\langle A_2^* y, y\rangle \ge \mu|y|^2 \quad \text{for all } y \in D(A_2^*). \tag{3.10}$$

Inequality (3.10) is similar to a condition assumed in [15].

Remark 3.2. Condition (3.9) is satisfied for instance when $\sigma : H \times H \times \Lambda \to \mathcal{L}(K, H)$ is bounded and Q is trace class in K. We also remark that since B_1 and B_2 are compact, we always have

$$\lim_{N\to\infty} \sup_{a\in\Lambda} \mathrm{Tr}\,(C(x, y, a)B_1Q_{1,N}) = 0 \tag{3.11}$$

and

$$\lim_{N\to\infty} \sup_{a\in\Lambda} \mathrm{Tr}\,(E(y, a)B_2Q_{2,N}) = 0, \tag{3.12}$$

see [14], page 237. In fact (3.11) is also a consequence of (3.9).

In the terminology of [11, 14], (3.3) means that A_1 satisfies the strong B_1-condition and A_2 satisfies the strong B_2-condition. If A is self-adjoint then A satisfies the strong B-condition with $B = (I + A)^{-1}$ and $c = 1$ and if in addition A is invertible then it satisfies the strong B-condition with $B = A^{-1}$ and $c = 0$. The strong B condition is generally satisfied by operators coming from elliptic equations. For instance, following [11, 14], let $\mathcal{O}$ be a bounded, smooth domain in $\mathbb{R}^n$ and

$$\begin{cases} Af := -\sum_{i,j}^n \partial_i(a_{ij}\partial_j f) + \sum_i^n b_i\partial_i f + ef \\ D(A) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O}), \end{cases} \tag{3.13}$$

where $a_{ij} = a_{ji}, b_i \in W^{1,\infty}(\mathcal{O}), e \in L^\infty(\mathcal{O})$ for $i, j \in \{1, \dots, n\}$, and there exists $\theta > 0$ such that

$$\sum_{i,j}^n a_{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2 \quad \forall x \in \mathcal{O}, \xi \in \mathbb{R}^n. \tag{3.14}$$

We assume that $e > 0$ is large enough so that A is maximal monotone. If $A_0 f = -\sum_{i,j}^n \partial_i(a_{ij}\partial_j f)$ with $D(A_0) = D(A)$, then A is self-adjoint and invertible so A_0 satisfies the strong B_0 condition with $B_0 = A_0^{-1}$ and $c = 0$. By elliptic regularity theory, B_0 is compact as an operator from $L^2(\mathcal{O})$ to $H_0^1(\mathcal{O})$ and thus $(A - A_0)^* B_0$ is compact in $L^2(\mathcal{O})$. Thus, by Example 3.15 of [14], the strong B-condition is satisfied for A with $B = \lambda B_0$ for some sufficiently big $\lambda > 0$ and for some c. We also remark that using a version of the Sobolevskii inequality (see, for instance, Thm. 1.1 of [21]) we can take $B_0 = \lambda_1(-\Delta)^{-1}, D(B_0) = D(A)$, for some $\lambda_1 > 0$ and then we can repeat the same arguments to obtain that A satisfies the strong B-condition with $B = \lambda_2(-\Delta)^{-1}$ for some $\lambda_2, c > 0$. Condition (3.2) requires some coerciveness of operators. It is always satisfied if A_1 is self-adjoint and invertible. It is also satisfied for the operator A in (3.13) if e is positive enough. However it cannot be satisfied by skew-adjoint operators. We refer the readers to [11] and [14], Section 3.1.1 for more on the strong B-condition.

Condition (3.4) is a strong monotonicity condition on the operator A_2 and the coefficients of the fast variable equation. It implies uniform exponential decay of second moments of the difference of two solutions of the fast variable equation (5.1) which allow to show uniform Lipschitz continuity estimates for viscosity solutions w_δ of discounted infinite horizon problems (5.2). They are also used to prove continuous dependence estimates for solutions of (5.1) which are essential in establishing the properties of the effective Hamiltonian $\overline{F}$ in Section 6. Condition (3.9) is a rather strong condition about uniform degeneracy of the noise in slow equation as the supremum there is taken over $a \in \Lambda$ and $y \in H$. However it is always satisfied if $h : H \times \Lambda \to \mathcal{L}(K, H)$ is bounded and Q is trace class, that is if the Q-Wiener process W_Q has nuclear covariance operator Q.

We recall that if Assumption 3.1 is satisfied then equation (1.1) has a unique mild solution, see [12] or [14], Section 1.4.2.

4. HJB equations and viscosity solutions

In this section we introduce various notions of viscosity solutions and prove a theorem about existence of a unique viscosity solution of equation (1.2). The definitions are variations of the general definitions of viscosity solutions from Chapter 3 of [14]. We want to emphasize that, contrary to equations in finite dimensional spaces, there is no single definition of a viscosity solution in infinite dimensional Hilbert spaces as equations in such spaces have different terms containing unbounded operators of various types. Thus each equation may require the use of different test functions and a special interpretation of such unbounded terms. This is why we will use several definitions of viscosity solution here. All definitions are related and use similar principles, however they are crafted to fit particular equations. Moreover we will also need a notion of a discontinuous viscosity solution (Def. 4.7) which requires a special definition. To avoid any confusion, each definition is stated explicitly.

Definition 4.1. A function $\psi : (0,T) \times H \times H \to \mathbb{R}$ is called a test function for equation (1.2) if $\psi(t,x,y) = \eta_1(t)\theta_1(|x|) + \eta_2(t)\theta_2(|y|) + \varphi(t,x,y)$, where:

(i) $\eta_1, \eta_2 \in C^1((0,T))$ and are strictly positive on $(0,T)$.
(ii) $\theta_1, \theta_2 \in C^2(\mathbb{R})$ are even, $\theta_1'(r) > 0, \theta_2'(r) \geq 0$ for $r \in (0,+\infty), \theta_1''(0) > 0$.
(iii) The function $\varphi \in C^{1,2,2}((0,T) \times H \times H)$ is weakly sequentially continuous and $\partial_t\varphi, A_1^* D_x\varphi,\ A_2^* D_y\varphi, D\varphi,$ $D^2\varphi$ are uniformly continuous on $(0,T) \times H \times H$.

Definition 4.2. A function $u : (0,T) \times H \times H \to \mathbb{R}$ is a viscosity subsolution of (1.2) if it is bounded from above, weakly sequentially upper semicontinuous and if for every test function ψ, whenever $u - \psi$ has a global maximum at (t,x,y), then $x \in H_{1,1}$ and

$$\begin{aligned}
&\partial_t\psi(t,x,y) - \lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|_{1,1}^2 - \frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y| - \langle x, A_1^* D_x\varphi(t,x,y)\rangle - \frac{1}{\varepsilon}\langle y, A_2^* D_y\varphi(t,x,y)\rangle \\
&- F\left(x, y, D_x\psi(t,x,y), \frac{1}{\varepsilon}D_y\psi(t,x,y), D_x^2\psi(t,x,y), \frac{1}{\varepsilon}D_y^2\psi(t,x,y), \frac{1}{\sqrt{\varepsilon}}D_{xy}^2\psi(t,x,y)\right) \geq 0.
\end{aligned} \tag{4.1}$$

A function $u : (0,T) \times H \times H \to \mathbb{R}$ is a viscosity supersolution of (1.2) is it is bounded from below, weakly sequentially lower semicontinuous and if for every test function ψ, whenever $u + \psi$ has a global minimum at (t,x,y), then $x \in H_{1,1}$ and

$$\begin{aligned}
&- \partial_t\psi(t,x,y) + \lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|_{1,1}^2 + \frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y| + \langle x, A_1^* D_x\varphi(t,x,y)\rangle + \frac{1}{\varepsilon}\langle y, A_2^* D_y\varphi(t,x,y)\rangle \\
&- F\left(x, y, -D_x\psi(t,x,y), -\frac{1}{\varepsilon}D_y\psi(t,x,y), -D_x^2\psi(t,x,y), -\frac{1}{\varepsilon}D_y^2\psi(t,x,y), -\frac{1}{\sqrt{\varepsilon}}D_{xy}^2\psi(t,x,y)\right) \leq 0.
\end{aligned} \tag{4.2}$$

A function u is a viscosity solution of (1.2) if it is a viscosity subsolution of (1.2) and a viscosity supersolution of (1.2).

Compared to the generic definition of a viscosity solution (see Def. 3.35 of [14]), we use a smaller class of test functions and we have two additional terms $\pm\lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|_{1,1}^2 \pm \frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y|$ in inequalities (4.1) and (4.2). These terms, which come from the radial parts of test functions, would have to be dropped according to Definition 3.35 of [14], in which nothing more is assumed about the unbounded operators besides their maximal monotonicity. This however would result in a loss of valuable coercive terms which make the viscosity solution techniques of this paper possible. The terms $\pm\lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|_{1,1}^2$ are the same as in the definitions in [19] and [14], Definition 3.90, and utilize the coercive condition (3.2). Without these terms in Definition 4.2 (and in Defs. 4.4 and 4.7), we would not be able to use the crucial technique of half-relaxed limits here. Note that, if $x \in D(A_1^*)$, then $\langle x, A_1^* D_x(\eta_1(t)\theta_1(|x|))\rangle = \eta_1(t)\frac{\theta_1(|x|)}{|x|}\langle x, A_1^* x\rangle \geq \lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|_{1,1}^2$. The terms $\pm\frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y|$ utilize the weaker coercive condition (3.10) and are important when dealing with the unboundedness of the fast variable y in the equations. Note that, if $y \in D(A_2^*)$, then $\frac{1}{\varepsilon}\langle y, A_2^* D_y(\eta_2(t)\theta_2(|y|))\rangle = \frac{1}{\varepsilon}\eta_2(t)\frac{\theta_2'(|y|)}{|y|}\langle y, A_2^* y\rangle \geq \frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y|$. Thus, due to the difference in the coercive conditions (3.2) and (3.10), Definition 4.2 is not symmetric with respect to the variables x and y.

We will also be dealing with HJB equations

$$\begin{cases} \partial_t u - \langle A_1 x, Du\rangle + G\left(x, Du, D^2 u\right) = 0 & \text{in } (0,T) \times H, \\ u(T,x) = g(x) & \text{for } x \in H \end{cases} \tag{4.3}$$

and

$$\delta v + \langle A_2 y, Dv\rangle + G_1\left(y, Dv, D^2 v\right) = 0 \quad \text{in } H, \tag{4.4}$$

where $\delta \geq 0$.

Definition 4.3. A function $\psi : (0,T)\times H \to \mathbb{R}$ is called a test function for equation (4.3) if $\psi(t,x) = \eta(t)\theta(|x|) + \varphi(t,x)$, where:

- (i) $\eta \in C^1((0,T))$ and is strictly positive on $(0,T)$.
- (ii) $\theta \in C^2(\mathbb{R})$ is even, $\theta'(r) > 0$ for $r \in (0,+\infty), \theta''(0) > 0$.
- (iii) The function $\varphi \in C^{1,2}((0,T)\times H)$ is weakly sequentially continuous and $\partial_t\varphi, A_1^* D\varphi, D\varphi, D^2\varphi$ are uniformly continuous on $(0,T)\times H$.

Definition 4.4. A function $u : (0,T)\times H \to \mathbb{R}$ is a viscosity subsolution of (4.3) is it is bounded from above, weakly sequentially upper semicontinuous and if for every test function ψ, whenever $u - \psi$ has a global maximum at (t,x), then $x \in H_{1,1}$, and

$$\partial_t\psi(t,x) - \lambda\eta(t)\frac{\theta'(|x|)}{|x|}|x|_{1,1}^2 - \langle x, A_1^* D\varphi(t,x)\rangle + G\left(x, D\psi(t,x), D^2\psi(t,x)\right) \geq 0. \tag{4.5}$$

A function $u : (0,T)\times H \to \mathbb{R}$ is a viscosity supersolution of (4.3) is it is bounded from below, weakly sequentially lower semicontinuous and if for every test function ψ, whenever $u + \psi$ has a global minimum at (t,x), then $x \in H_{1,1}$, and

$$-\partial_t\psi(t,x) + \lambda\eta(t)\frac{\theta'(|x|)}{|x|}|x|_{1,1}^2 + \langle x, A_1^* D\varphi(t,x)\rangle + G\left(x, -D\psi(t,x), -D^2\psi(t,x)\right) \leq 0. \tag{4.6}$$

A function u is a viscosity solution of (4.3) if it is a viscosity subsolution of (4.3) and a viscosity supersolution of (4.3).

Definition 4.5. A function $\psi : H \to \mathbb{R}$ is called a test function for equation (5.7) if $\psi(y) = \theta(|y|) + \varphi(y)$, where $\theta \in C^2(\mathbb{R})$ is even, $\theta'(r) \geq 0$ for $r \in (0,+\infty)$, $\varphi \in C^2(H)$ is weakly sequentially continuous and $A_2^* D\varphi, D\varphi, D^2\varphi$ are uniformly continuous on H.

Definition 4.6. A function $v : H \to \mathbb{R}$ is a viscosity subsolution of (5.7) is it is bounded from above, weakly sequentially upper semicontinuous and if for every test function ψ, whenever $v - \psi$ has a global maximum at y, then

$$\delta v(y) + \mu\theta'(|y|)|y| + \langle y, A_2^* D\varphi(y)\rangle + G_1(y, D\psi(y), D^2\psi(y)) \leq 0. \tag{4.7}$$

A function $v : H \to \mathbb{R}$ is a viscosity supersolution of (5.7) is it is bounded from below, weakly sequentially lower semicontinuous and if for every test function ψ, whenever $v + \psi$ has a global minimum at y, then

$$\delta v(y) - \mu\theta'(|y|)|y| - \langle y, A_2^* D\varphi(y)\rangle + G_1(y, -D\psi(y), -D^2\psi(y)) \geq 0. \tag{4.8}$$

A function u is a viscosity solution of (5.7) if it is a viscosity subsolution of (5.7) and a viscosity supersolution of (5.7).

To use the technique of half-relaxed limits we will need a notion of a discontinuous viscosity solution. For a function $v : (0,T)\times H \to \mathbb{R}$ we define its upper and lower semicontinuous envelopes in $\mathbb{R}\times H_{1,-1}$ by

$$v^*(t,x) = \limsup\{v(s,y) : |s-t| \to 0, |y-x|_{1,-1} \to 0\},$$

$$v_*(t,x) = \liminf\{v(s,y) : |s-t| \to 0, |y-x|_{1,-1} \to 0\}.$$

Definition 4.7. A function $u : (0,T) \times H \to \mathbb{R}$ is a discontinuous viscosity subsolution of (4.3) is it is bounded from above and if for every test function ψ from Definition 4.3 satisfying in addition $\theta(r) \to +\infty$ as $r \to +\infty$, whenever $(u - \eta(\cdot)\theta(|\cdot|))^* - \varphi$ has a global maximum at (t,x), then $x \in H_{1,1}$, and

$$\partial_t \psi(t,x) - \lambda\eta(t)\frac{\theta'(|x|)}{|x|}|x|^2_{1,1} - \langle x, A_1^* D\varphi(t,x)\rangle + G\left(x, D\psi(t,x), D^2\psi(t,x)\right) \geq 0. \tag{4.9}$$

A function $u : (0,T) \times H \to \mathbb{R}$ is a discontinuous viscosity supersolution of (4.3) is it is bounded from below and if for every test function ψ from Definition 4.3 satisfying in addition $\theta(r) \to +\infty$ as $r \to +\infty$, whenever $(u + \eta(\cdot)\theta(|\cdot|))_* + \varphi$ has a global minimum at (t,x), then $x \in H_{1,1}$, and

$$-\partial_t \psi(t,x) + \lambda\eta(t)\frac{\theta'(|x|)}{|x|}|x|^2_{1,1} + \langle x, A_1^* D\varphi(t,x)\rangle + G\left(x, -D\psi(t,x), -D^2\psi(t,x)\right) \leq 0. \tag{4.10}$$

A function u is a discontinuous viscosity solution of (4.3) if it is a discontinuous viscosity subsolution of (4.3) and a discontinuous viscosity supersolution of (4.3).

We point out that even if u above is continuous but not weakly sequentially continuous it still falls into a category of a discontinuous viscosity solution.

We note that for test functions θ_1, θ_2 from Definition 4.2, we have $\theta_i'(0) = 0$ and the function $\tilde{\theta}_i(x) = \theta_i(|x|) \in C^2(H), i = 1,2$. Moreover, since $\theta_1'(r)/r \to \theta_1''(0) > 0$ as $r \to 0$, the term $\theta_1'(|x|)/|x|$ is bounded away from 0 on bounded sets and, after modifying $\theta_1(r)$ for large r we can always assume that $\theta_1'(|x|)/|x| \geq c > 0$ on H. We also remark that we could have assumed that the maxima/minima in all definitions are local, and we can always assume that the maxima/minima in Definitions 4.2, 4.4, 4.6 are strict and the maxima/minima in Definition 4.7 are strict in the $|\cdot| \times |\cdot|_{1,-1}$ norm (see [19], Lem. 3.6 and [14], Lem. 3.37). Moreover, since we only deal with bounded viscosity sub/super-solutions and φ has at most quadratic growth, we can always assume that $\theta_i(|x|) \leq C(1+|x|^2), i = 1,2, x \in H$. We say that a function f has a strict maximum at (t,x,y) if it has a maximum at (t,x,y) and if whenever $f(t_n,x_n,y_n) \to f(t,x,y)$ then (t_n,x_n,y_n) converges (in a proper norm) to (t,x,y).

Theorem 4.8. *Let Assumption 3.1 be satisfied. Then, for every $\varepsilon > 0$, there exists a unique bounded viscosity solution $u_\varepsilon \in C((0,T] \times H \times H)$ in the sense of Definition 4.2 of equation (1.2) satisfying*

$$\lim_{t \to T, t<T} \sup_{x,y \in H} |(u_\varepsilon(t,x,y) - f(e^{-(T-t)A_1}x))| = 0. \tag{4.11}$$

Proof. The existence of a unique viscosity solution in the sense of Definition 3.35 of [14] given by the value function of the associated optimal control problem follows from Theorem 3.67 of [14] after we notice that the operator $\tilde{A}(x,y) := (A_1 x, \frac{1}{\varepsilon}A_2 y)$ is a maximal monotone operator in $H \times H$, $\tilde{B}(x,y) := (B_1 x, B_2 y)$ is such that $\tilde{B} \in S(H \times H), B > 0$ and

$$\langle (\tilde{A}^*\tilde{B} + c\tilde{B})(x,y),(x,y)\rangle \geq |x|^2 + |y|^2 \quad \text{for some } c \geq 0 \text{ and all } x,y \in H,$$

that is $\tilde{A}$ satisfies the strong $\tilde{B}$-condition.

Uniqueness of viscosity solutions of (1.2) is a consequence of a comparison principle whose proof follows the proof of Theorem 3.54 of [14]. In the proof of comparison the terms $\lambda\eta_1(t)\frac{\theta_1'(|x|)}{|x|}|x|^2_{1,1}$ and $\frac{\mu}{\varepsilon}\eta_2(t)\theta_2'(|y|)|y|$ can be dropped so that we can work with the definition of viscosity solution from [14]. The radial parts of test functions used in [14] were more general to accommodate unbounded solutions with possible exponential growth however, since we deal with bounded subsolutions and supersolutions here, the proof of comparison principle (Thm. 3.54

of [14]) is much simpler and works after standard modifications if we replace the radial test functions used there with test functions $\delta e^{K(T-t)}(|x|^2+|y|^2)$. The full proof of comparison principle for a similar equation (7.1) in $(0,T]\times H$ is also given in the proof of Theorem 7.1. We can basically follow this proof here after adjusting for the extra variable. We also note that we do not have to use perturbed optimization in this paper (like in the proof of Theorem 3.54 of [14]) as our operators B_1, B_2 are compact and hence the functions in the proof are weakly sequentially upper semicontinuous. We leave the details to the readers. It remains to explain a few details of the proof that the value function is also a solution in the sense of Definition 4.2 and satisfies (4.11).

Following the proof of Lemma 3.23 of [14] we easily obtain that for every $a(\cdot)\in\mathcal{U}_t$, if $X^i_\varepsilon, Y^i_\varepsilon, i=1,2$, are the unique mild solutions of (1.1) with initial conditions $X^i_\varepsilon(t)=x_i, Y^i_\varepsilon(t)=y_i$, then there is a constant $C_\varepsilon(T)$, independent of $a(\cdot)$ and $x_i, y_i\in H$, such that

$$\begin{aligned}&\sup_{s\in[t,T]}\left(\mathbb{E}[|X^1_\varepsilon(s)-X^2_\varepsilon(s)|^2_{1,-1}+|Y^1_\varepsilon(s)-Y^2_\varepsilon(s)|^2_{2,-1}]\right)\\&+\int_t^T\mathbb{E}[|X^1_\varepsilon(s)-X^2_\varepsilon(s)|^2+|Y^1_\varepsilon(s)-Y^2_\varepsilon(s)|^2]\mathrm{d}s\le C_\varepsilon(T)(|x_1-x_2|^2_{1,-1}+|y_1-y_2|^2_{2,-1})\end{aligned}\tag{4.12}$$

and

$$\mathbb{E}[|X^1_\varepsilon(s)-X^2_\varepsilon(s)|^2]\le C_\varepsilon(T)|y_1-y_2|^2_{2,-1}+\frac{C_\varepsilon(T)}{s-t}|x_1-x_2|^2_{1,-1}\quad\text{for all } s\in(t,T].\tag{4.13}$$

Then, arguing like in the proof of Proposition 3.62 of [14], we obtain that there exist moduli ρ,ρ_τ such that, if u_ε is the value function associated to equation (1.2), we have for all $0\le t\le T-\tau, x_i, y_i\in H, i=1,2$,

$$|u_\varepsilon(t,x_1,y_1)-u_\varepsilon(t,x_2,y_2)|\le\rho(|y_1-y_2|_{2,-1})+\rho_\tau(|x_1-x_2|_{1,-1}),$$

the dynamic programming principle is satisfied, $u_\varepsilon\in C((0,T]\times H\times H)$ and there exists a modulus $\tilde\rho$ such that for every $0\le t\le T, x,y\in H$,

$$|u_\varepsilon(t,x,y)-f(e^{-(T-t)A_1}x)|\le\tilde\rho(T-t).$$

Using the dynamic programming principle one shows as in the proof of Theorem 3.66 of [14] that the value function u_ε is a viscosity solution of (1.2) in the sense of Definition 3.35 of [14]. It remains to explain that the x components of the points where the maxima and minima occur are in $H_{1,1}$ and we have the extra terms (which are dropped in Definition 3.35 of [14]) in inequalities (4.1) and (4.2).

To simplify the notation we define $\tilde b(x,y,a)=(b(x,y,a),\frac{1}{\varepsilon}g(x,y,a))$, $\tilde\sigma(x,y,a)w=(\sigma(x,y,a)w, \frac{1}{\sqrt{\varepsilon}}h(y,a)w), w\in H$. We notice that

$$\langle\tilde A^*(x,y),(x,y)\rangle\ge\lambda|x|^2_{1,1}+\frac{\mu}{\varepsilon}|y|^2\quad\text{for all }(x,y)\in D(\tilde A^*).$$

Moreover, the function $\tilde b$ satisfies (3.5) and $\tilde\sigma$ satisfies (3.6). Then, for $a(\cdot)\in\mathcal{U}^\mu_t$, equation (1.1) can be rewritten for the variable $Z(t)=(X_\varepsilon(t),Y_\varepsilon(t))$ as

$$\begin{cases}\mathrm{d}Z(s)=\left[-\tilde AZ(s)+\tilde b(Z(s),a(s))\right]\mathrm{d}s+\tilde\sigma(Z(s),a(s))\mathrm{d}W_Q(s)\\ Z(t)=z:=(x,y)\in H\times H.\end{cases}\tag{4.14}$$

In the rest of the proof we will use $|\cdot|$ to denote both the norm in H and the norm in $H\times H$. Equation (4.14) has a unique mild solution which satisfies the estimates of Theorem 1.130 of [14], in particular

$$\mathbb{E}[\sup_{t\le s\le T}|Z(s)|^p]\le C_p(T)(1+|z|^p) \tag{4.15}$$

for every $p\ge 2$.

We set for $N\ge 1$, $P_N z=(P_{1,N}x,P_{2,N}y)$,

$$\tilde{A}_N z=(P_N\tilde{A}^*P_N)^*z=((P_{1,N}A_1^*P_{1,N})^*x,(P_{2,N}A_2^*P_{2,N})^*y)=:(A_{1,N}x,A_{2,N}y)$$

and define $Z_N=(X_N,Y_N)$ to be the solution of

$$\begin{cases} \mathrm{d}Z_N(s)=\left[-\tilde{A}_N Z_N(s)+P_N\tilde{b}(Z_N(s),a(s))\right]\mathrm{d}s+P_N\tilde{\sigma}(Z_N(s),a(s))\mathrm{d}W_Q(s)\\ Z_N(t)=P_N(x,y). \end{cases}$$

Recall that $\tilde{A}_N$ is a bounded operator such that $-\tilde{A}_N$ generates a semigroup of contractions and, by Lemma 3.83 of [14], for every z and $T>0$,

$$e^{-t\tilde{A}_N}z\to e^{-t\tilde{A}}z\quad \text{uniformly on } [0,T]. \tag{4.16}$$

We have

$$\begin{aligned} Z(s)-Z_N(s)&=e^{-(s-t)\tilde{A}}z-e^{-(s-t)\tilde{A}_N}P_Nz+\int_t^s(e^{-(s-r)\tilde{A}}-e^{-(s-r)\tilde{A}_N})\tilde{b}(Z(r),a(r))\mathrm{d}r\\ &+\int_t^s(e^{-(s-r)\tilde{A}}-e^{-(s-r)\tilde{A}_N})\tilde{\sigma}(Z(r),a(r))\mathrm{d}W_Q(r)\\ &+\int_t^s e^{-(s-r)\tilde{A}_N}Q_N\tilde{b}(Z(r),a(r))\mathrm{d}r+\int_t^s e^{-(s-r)\tilde{A}_N}Q_N\tilde{\sigma}(Z(r),a(r))\mathrm{d}W_Q(r)\\ &+\int_t^s e^{-(s-r)\tilde{A}_N}P_N(\tilde{b}(Z(r),a(r))-\tilde{b}(Z_N(r),a(r)))\mathrm{d}r\\ &+\int_t^s e^{-(s-r)\tilde{A}_N}P_N(\tilde{\sigma}(Z(r),a(r))-\tilde{\sigma}(Z_N(r),a(r)))\mathrm{d}W_Q(r). \end{aligned}$$

Therefore

$$\begin{aligned} &\mathbb{E}|Z(s)-Z_N(s)|^2\le C|e^{-(s-t)\tilde{A}}z-e^{-(s-t)\tilde{A}_N}P_Nz|^2+C\int_t^T\mathbb{E}|(e^{-(s-r)\tilde{A}}-e^{-(s-r)\tilde{A}_N})\tilde{b}(Z(r),a(r))|^2\mathrm{d}r\\ &+C\int_t^T\mathbb{E}\big[\mathrm{Tr}\big((e^{-(s-r)\tilde{A}}-e^{-(s-r)\tilde{A}_N})\tilde{\sigma}(Z(r),a(r))Q^{\frac12}\big)\big((e^{-(s-r)\tilde{A}}-e^{-(s-r)\tilde{A}_N})\tilde{\sigma}(Z(r),a(r))Q^{\frac12}\big)^*\big]\mathrm{d}r\\ &+C\int_t^T\mathbb{E}|Q_N\tilde{b}(Z(r),a(r))|^2\mathrm{d}r+C\int_t^T\mathbb{E}\big[\mathrm{Tr}\big(Q_N\tilde{\sigma}(Z(r),a(r))Q^{\frac12}\big)\big(Q_N\tilde{\sigma}(Z(r),a(r))Q^{\frac12}\big)^*\big]\mathrm{d}r\\ &+C\int_t^s\mathbb{E}|Z(r)-Z_N(r)|^2\mathrm{d}r. \end{aligned} \tag{4.17}$$

We now use (4.16) to estimate the first term in the right hand side of (4.17), (4.16) and the dominated convergence theorem to estimate the second and third terms, and the dominated convergence theorem to estimate the fourth and fifth terms to obtain

$$\mathbb{E}|Z(s)-Z_N(s)|^2 \le a_N + C\int_t^s \mathbb{E}|Z(r)-Z_N(r)|^2 \mathrm{d}r,$$

where $a_N \to 0$ as $N\to\infty$. Therefore, by Gronwall's inequality, it follows that

$$\mathbb{E}|Z(s)-Z_N(s)|^2 \le a_N e^{C(T-t)} \quad \text{for all } s\in[t,T]. \tag{4.18}$$

Let now $\eta_1(s)\theta_1(|x|)+\eta_2(s)\theta_2(|y|)+\varphi(s,x,y)$ be a test function. We can assume that the functions $\tilde\theta_1(x)=\theta_1(|x|), \tilde\theta_2(y)=\theta_1(|y|)$ have uniformly continuous first and second Fréchet derivatives, bounded second Fréchet derivatives on H and $0<c\le\theta_1'(r)/r$ is uniformly continuous and bounded on $(0,+\infty)$. Thus, if $\varepsilon>0$, we have by Itô's formula

$$\begin{aligned}
&\mathbb{E}\eta_1(s)\theta_1(|X_N(s)|)+\mathbb{E}\eta_2(s)\theta_2(|Y_N(s)|) = \eta_1(t)\theta_1(|P_{1,N}x|)+\eta_2(t)\theta_2(|P_{2,N}y|)\\
&+\int_t^s \mathbb{E}\eta_1'(r)\theta_1(|X_N(s)|)\mathrm{d}r + \int_t^s \mathbb{E}\eta_2'(r)\theta_2(|Y_N(s)|)\mathrm{d}r\\
&-\int_t^s \mathbb{E}\eta_1(r)\frac{\theta_1'(|X_N(r)|)}{|X_N(r)|}\langle A_{1,N}X_N(r),X_N(r)\rangle\mathrm{d}r - \int_t^s \mathbb{E}\frac{1}{\varepsilon}\eta_2(r)\frac{\theta_2'(|Y_N(r)|)}{|Y_N(r)|}\langle A_{2,N}Y_N(r),Y_N(r)\rangle\mathrm{d}r\\
&+\int_t^s \mathbb{E}\eta_1(r)\langle D\tilde\theta_1(X_N(r)), b(X_N(r),Y_N(r),a(r))\rangle\mathrm{d}r\\
&+\int_t^s \mathbb{E}\frac{1}{\varepsilon}\eta_2(r)\langle D\tilde\theta_2(Y_N(r)), g(X_N(r),Y_N(r),a(r))\rangle\mathrm{d}r\\
&+\frac12\int_t^s \mathbb{E}\eta_1(r)\mathrm{Tr}[(P_{1,N}\sigma(X_N(r),Y_N(r),a(r))Q^{\frac12})(P_{1,N}\sigma(X_N(r),Y_N(r),a(r))Q^{\frac12})^*D^2\tilde\theta_1(X_N(r))]\mathrm{d}r\\
&+\frac12\int_t^s \mathbb{E}\frac{1}{\varepsilon}\eta_2(r)\mathrm{Tr}[(P_{2,N}h(Y_N(r),a(r))Q^{\frac12})(P_{2,N}h(Y_N(r),a(r))Q^{\frac12})^*D^2\tilde\theta_2(Y_N(r))]\mathrm{d}r.
\end{aligned} \tag{4.19}$$

We now estimate

$$\langle A_{1,N}X_N(r),X_N(r)\rangle \ge \lambda|X_N(r)|_{1,1}^2, \quad \langle A_{2,N}Y_N(r),Y_N(r)\rangle \ge \mu|Y_N(r)|^2. \tag{4.20}$$

Then (4.19), together with (4.15) and (4.20), yields

$$\int_t^s \mathbb{E}[\eta_1(r)\frac{\theta_1'(|X_N(r)|)}{|X_N(r)|}|X_N(r)|_{1,1}^2]\mathrm{d}r < C \quad \text{for all } N\ge 1. \tag{4.21}$$

Denote $W_N(r)=\eta_1(r)\frac{\theta_1'(|X_N(r)|)}{|X_N(r)|}, W_\varepsilon(r)=\eta_1(r)\frac{\theta_1'(|X_\varepsilon(r)|)}{|X_\varepsilon(r)|}$. Inequality (4.21) implies that, up to a subsequence, $\sqrt{W_N}X_N \rightharpoonup \tilde X$ in $M_\mu^2(t,s;H_{1,1})$. Since, by (4.18), $\sqrt{W_N}X_N\to\sqrt{W_\varepsilon}X_\varepsilon$ in $M_\mu^2(t,s;H)$, we thus obtain $\sqrt{W_\varepsilon}X_\varepsilon = \tilde X\in M_\mu^2(t,s;H_{1,1})$.

We can now pass to the limit as $N\to\infty$ in (4.19) using the continuity properties of the coefficients and $\tilde\theta_1$ and its derivatives, lower semicontinuity of the norm with respect to weak convergence, (4.15), (4.18), (4.20) and the dominated convergence theorem, to obtain

$$
\begin{aligned}
&\mathbb{E}\eta_1(s)\theta_1(|X_\varepsilon(s)|)+\mathbb{E}\eta_2(s)\theta_2(|Y_\varepsilon(s)|)\le\eta_1(t)\theta_1(|x|)+\eta_2(t)\theta_2(|y|)\\
&+\int_t^s\mathbb{E}\eta_1'(r)\theta_1(|X_\varepsilon(r)|)\mathrm{d}r+\int_t^s\mathbb{E}\eta_2'(r)\theta_2(|Y_\varepsilon(r)|)\mathrm{d}r\\
&-\int_t^s\mathbb{E}\lambda\eta_1(r)\frac{\theta_1'(|X_\varepsilon(r)|)}{|X_\varepsilon(r)|}|X_\varepsilon(r)|_{1,1}^2\mathrm{d}r-\int_t^s\mathbb{E}\frac{\mu}{\varepsilon}\eta_2(r)\theta_2'(|Y_\varepsilon(r)|)|Y_\varepsilon(r)|\mathrm{d}r\\
&+\int_t^s\mathbb{E}\eta_1(r)\langle D\tilde\theta_1(X_\varepsilon(r)),b(X_\varepsilon(r),Y_\varepsilon(r),a(r))\rangle\mathrm{d}r+\int_t^s\mathbb{E}\frac{1}{\varepsilon}\eta_2(r)\langle D\tilde\theta_2(Y_\varepsilon(r)),g(X_\varepsilon(r),Y_\varepsilon(r),a(r))\rangle\mathrm{d}r\\
&+\int_t^s\mathbb{E}\eta_1(r)\mathrm{Tr}[C(X_\varepsilon(r),Y_\varepsilon(r),a(r))D^2\tilde\theta_1(X_\varepsilon(r))]\mathrm{d}r+\int_t^s\mathbb{E}\frac{1}{\varepsilon}\eta_2(r)\mathrm{Tr}[E(Y_\varepsilon(r),a(r))D^2\tilde\theta_2(Y_\varepsilon(r))]\mathrm{d}r.
\end{aligned}
\tag{4.22}
$$

Itô's formula for test function φ is proved in Proposition 1.165 and Lemma 3.65 of [14].

We now briefly sketch the proof that viscosity solution inequalities are satisfied. We will only do it for the subsolution part. We follow the proof of Theorem 3.66 of [14]. Suppose that $u_\varepsilon-\psi$ has a global maximum at (t,x,y) for a test function ψ. Recall that u_ε is the value function. We take $a\in\Lambda$ and define a constant control $a(\cdot)=a$ on any reference probability space μ. Using the dynamic programming principle, (4.22) and Itô's formula for φ, we obtain for $s_n=t+\frac{1}{n}$

$$
\begin{aligned}
0\le&-\frac{1}{n}\mathbb{E}\bigg[\int_t^{s_n}\lambda\eta_1(r)\frac{\theta_1'(|X_\varepsilon(r)|)}{|X_\varepsilon(r)|}|X_\varepsilon(r)|_{1,1}^2\mathrm{d}r\bigg]\\
&+\frac{1}{n}\mathbb{E}\bigg[\int_t^{s_n}\bigg(\partial_t\psi(r,X_\varepsilon(r),Y_\varepsilon(r))-\frac{\mu}{\varepsilon}\eta_2(r)\theta_2'(|Y_\varepsilon(r)|)|Y_\varepsilon(r)|\\
&-\langle X_\varepsilon(r),A_1^*D_x\varphi(r,X_\varepsilon(r),Y_\varepsilon(r))\rangle-\frac{1}{\varepsilon}\langle Y_\varepsilon(r),A_2^*D_y\varphi(r,X_\varepsilon(r),Y_\varepsilon(r))\rangle\\
&+\langle D_x\psi(r,X_\varepsilon(r),Y_\varepsilon(r)),b(X_\varepsilon(r),Y_\varepsilon(r),a(r))\rangle+\frac{1}{\varepsilon}\langle D_y\psi(r,X_\varepsilon(r),Y_\varepsilon(r)),g(X_\varepsilon(r),Y_\varepsilon(r),a)\rangle\\
&+\mathrm{Tr}[C(X_\varepsilon(r),Y_\varepsilon(r),a)D_x^2\psi(r,X_\varepsilon(r),Y_\varepsilon(r))]+\frac{1}{\varepsilon}\mathrm{Tr}[E(X_\varepsilon(r),a)D_y^2\psi(r,X_\varepsilon(r),Y_\varepsilon(r))]\\
&+\frac{1}{\sqrt{\varepsilon}}\mathrm{Tr}[D(X_\varepsilon(r),Y_\varepsilon(r),a)D_{xy}^2\psi(r,X_\varepsilon(r),Y_\varepsilon(r))]+l(X_\varepsilon(r),Y_\varepsilon(r),a)\bigg)\mathrm{d}r\bigg].
\end{aligned}
\tag{4.23}
$$

Since, by Jensen's inequality (using previous notation),

$$
\mathbb{E}\left|\frac{1}{n}\int_t^{s_n}\sqrt{W_\varepsilon(r)}X_\varepsilon(r)\mathrm{d}r\right|_{1,1}^2\le\frac{1}{n}\mathbb{E}\int_t^{s_n}W_\varepsilon(r)|X_\varepsilon(r)|_{1,1}^2\mathrm{d}r,
\tag{4.24}
$$

using the properties of test functions and moment estimates for $X_\varepsilon,Y_\varepsilon$ (see Thm. 1.130 of [14]), we thus have that the variables

$$
U_n=\frac{1}{n}\int_t^{s_n}\sqrt{W_\varepsilon(r)}X_\varepsilon(r)\mathrm{d}r
$$

are bounded in $L^2(\Omega;H_{1,1})$ so, up to a subsequence, they converge weakly in this space to $U_n\rightharpoonup\tilde U$ in $L^2(\Omega;H_{1,1})$ for some element $\tilde U$ in this space. But, for instance by (1.38) of [14],

$$
U_n\to\sqrt{\eta_1(t)}\frac{\sqrt{\theta_1'(|x|)}}{\sqrt{|x|}}x\quad\text{in }L^2(\Omega;H)
\tag{4.25}
$$

so we get $x \in H_{1,1}$. We can now pass to the lim sup as $n \to \infty$ in (4.23) using (4.24), (4.25), the weak lower semicontinuity of the norm, the properties of test functions, and the moment and continuity estimates of Theorem 1.130 of [14] for $X_\varepsilon, Y_\varepsilon$. It then remains to take the infimum over all $a \in \Lambda$ in the obtained inequality. The proof of the supersolution property, which is explained in the proof of Theorem 3.66 of [14], is a little more complicated but it requires similar modifications. □

Remark 4.9. In order to prove Theorem 4.8 we can replace $|x_1 - x_2|_{1,-1}$ by $|x_1 - x_2|$ in (3.7) and only require $|l(x_1, y_1, a) - l(x_2, y_2, a)| \le \omega(|x_1 - x_2| + |y_1 - y_2|)$.

5. Discounted problems in the fast variable and effective Hamiltonian

This section is devoted to the analysis of discounted infinite horizon problems in the fast variable, the solvability of the ergodic control problem in the fast variable and the definition of the effective Hamiltonian. The overall approach follows [1, 4].

Theorem 5.1. *Let Assumption 3.1 be satisfied. Let $(x, p, X) \in H \times H \times S(H)$ be fixed. Define for $\delta > 0$*

$$w_\delta(y) = \inf_{a(\cdot)\in\mathcal{U}_0} \mathbb{E} \int_0^{+\infty} e^{-\delta s} \Big(\mathrm{Tr}(C(x, Y(s), a(s))X) + \langle b(x, Y(s), a(s)), p\rangle + l(x, Y(s), a(s)) \Big) \mathrm{d}s,$$

where $Y(s)$ is the solution of

$$\begin{cases} \mathrm{d}Y(s) = [-A_2 Y(s) + g(x, Y(s), a(s))]\, \mathrm{d}s + h(Y(s), a(s))\mathrm{d}W_Q(s) \\ Y(0) = y. \end{cases} \tag{5.1}$$

The function w_δ is the unique bounded viscosity solution of the equation

$$\delta w_\delta + \langle A_2 y, D_y w_\delta\rangle + F\left(x, y, p, D_y w_\delta, X, D_y^2 w_\delta, 0\right) = 0 \quad \text{in } H, \tag{5.2}$$

there exists a constant $C = C(p, X)$ such that for every $\delta > 0$

$$|w_\delta(y_1) - w_\delta(y_2)| \le C|y_1 - y_2| \quad \text{for all } y_1, y_2, x \in H \tag{5.3}$$

and δw_δ converges, as $\delta \to 0$, uniformly on bounded sets in H to a constant. We denote

$$\overline{F}(x, p, X) := -\lim_{\delta\to 0} \delta w_\delta.$$

Proof. The fact that w_δ is the unique bounded viscosity solution of equation (5.2) follows from Theorem 3.76 of [14] after a few modifications, using Itô's formula (4.22), similar to what we did in the proof of Theorem 4.8. Notice that, since all coefficient functions C, b, l, g, h here are bounded, in the proof of comparison principle, Theorem 3.58 of [14], it is enough to use our test functions.

Let

$$M = \sup_{y\in H} |F(x, y, p, 0, X, 0, 0)|. \tag{5.4}$$

Then $\overline{w} = \frac{M}{\delta}, \underline{w} = -\frac{M}{\delta}$ are respectively a viscosity supersolution and a viscosity subsolution of (5.2) so by comparison we obtain $|\delta w_\delta| \le M$. We will now prove (5.3).

Let $Y(s)$ be the mild solution of (5.1). For $N \ge 1$, let $\tilde{A}_{2,N} = (P_{2,N} A_2^* P_{2,N})^*$ and let Y_N be the solution of

$$\begin{cases} \mathrm{d}Y_N(s) = \Big[-\tilde{A}_{2,N} Y_N(s) + P_{2,N} g(x, Y_N(s), a(s))\Big]\,\mathrm{d}s + P_{2,N} h(Y_N(s), a(s))\mathrm{d}W_Q(s) \\ Y_N(0) = P_{2,N} y. \end{cases} \tag{5.5}$$

Repeating the arguments in the proof of Theorem 4.8 we obtain that for every $t > 0$ there is a sequence (a_N^t) such that

$$\mathbb{E}[|Y(s) - Y_N(s)|^2] \le a_N^t \quad \text{for } 0 \le s \le t,$$

where $a_N^t \to 0$ as $N \to +\infty$.

Let $Y_1(s), Y_2(s)$ be the mild solutions of (5.1) with initial conditions $Y_1(0) = y_1, Y_1(0) = y_2$ and let $Y_N^1(s), Y_N^2(s)$ be the solutions of (5.5) with initial conditions $Y_N^1(0) = P_{2,N} y_1, Y_N^2(0) = P_{2,N} y_2$. Applying Itô's formula and using (3.4) and the properties of the trace, we have

$$\begin{aligned} & e^{2\mu t}\mathbb{E}|Y_N^1(t) - Y_N^2(t)|^2 \le |y_1 - y_2|^2 + 2\mu \mathbb{E}\int_0^t e^{2\mu s}|Y_N^1(s) - Y_N^2(s)|^2 \mathrm{d}s \\ & - 2\mathbb{E}\int_0^t e^{2\mu s}\langle A_2^*(Y_N^1(s) - Y_N^2(s)), Y_N^1(s) - Y_N^2(s)\rangle \mathrm{d}s \\ & + 2\mathbb{E}\int_0^t e^{2\mu s}\langle g(x, Y_N^1(s), a(s)) - g(x, Y_N^2(s), a(s)), Y_N^1(s) - Y_N^2(s)\rangle \mathrm{d}s \\ & + \mathbb{E}\int_0^t e^{2\mu s}\mathrm{Tr}\big[(h(Y_N^1(s), a(s)) - h(Y_N^2(s), a(s)))Q^{\frac{1}{2}}) \\ & \qquad\qquad ((h(Y_N^1(s), a(s)) - h(Y_N^2(s), a(s)))Q^{\frac{1}{2}})^*\big]\mathrm{d}s \\ & \qquad\qquad\quad \le |y_1 - y_2|^2. \end{aligned}$$

Letting $N \to \infty$ above we conclude

$$\mathbb{E}|Y_1(t) - Y_2(t)|^2 \le e^{-2\mu t}|y_1 - y_2|^2.$$

It now follows from (3.6) that

$$\begin{aligned} |w_\delta(y_1) - w_\delta(y_2)| & \le \sup_{a(\cdot)\in\mathcal{U}_0} \mathbb{E}\int_0^{+\infty} e^{-\delta s} C(1 + |p| + \|X\|)|Y_1(s) - Y_2(s)|\mathrm{d}s \\ & \le \int_0^{+\infty} e^{-\mu s} C(1 + |p| + \|X\|)|y_1 - y_2|\mathrm{d}s \le C(1 + |p| + \|X\|)|y_1 - y_2|. \end{aligned}$$

We now choose a sequence δ_n such that $a_n = \delta_n w_{\delta_n}(0) \to \nu$ for some ν. Then, by (5.3), $\delta_n w_{\delta_n}(y) \to \nu$ uniformly on bounded sets of H. Suppose there is another sequence δ_n^1 such that $b_n = \delta_n^1 w_{\delta_n^1}(0) \to \nu_1$ for some ν_1 and then $\delta_n^1 w_{\delta_n^1}(y) \to \nu_1$ uniformly on bounded sets of H. Suppose without loss of generality that $\nu < \nu_1$. Denote

$$v_n(y) = w_{\delta_n}(y) + \kappa|y|^2 - w_{\delta_n}(0), \quad z_n(y) = w_{\delta_n^1}(y) - w_{\delta_n^1}(0) + 1,$$

where $\kappa > 0$ will be determined later. Using the definition of viscosity supersolution, (3.10) and the boundedness of g and h, we check that v_n is a viscosity supersolution of

$$\begin{aligned}\delta_n v_n + \langle A_2 y, Dv_n\rangle + F\left(x, y, p, Dv_n, X, D^2 v_n, 0\right)& \\ = -a_n + \kappa((2\mu + \delta_n)|y|^2 - C(|y| + 1)) &=: f_n(y).\end{aligned} \tag{5.6}$$

Moreover, z_n is a viscosity subsolution of

$$\delta_n^1 z_n + \langle A_2 y, Dz_n\rangle + F\left(x, y, p, Dz_n, X, D^2 z_n, 0\right) = -b_n + \delta_n^1$$

and hence z_n is a viscosity subsolution of

$$\begin{aligned}\delta_n z_n + \langle A_2 y, Dz_n\rangle + F\left(x, y, p, Dz_n, X, D^2 z_n, 0\right) &= -b_n + \delta_n^1 + (\delta_n - \delta_n^1) z_n(y) \\ &\le -b_n + C(\delta_n + \delta_n^1)(1 + |y|) =: f_n^1(y).\end{aligned}$$

We now observe that if n is large enough then

$$f_n(y) - f_n^1(y) \ge -a_n + b_n - C_1\kappa \to \nu_1 - \nu - C_1\kappa$$

for some absolute constant $C_1 > 0$ so if $\kappa = (\nu_1 - \nu)/(2C_1)$ then, for large n, we have $f_n > f_n^1$ on H. Thus, z_n is a viscosity subsolution of (5.6) and v_n is a viscosity supersolution of (5.6) and hence, by comparison principle, $z_n \le v_n$ on H. But this is impossible since $z_n(0) = 1 > 0 = v_n(0)$. This proves that we must have $\delta w_\delta(0) \to \nu$ as $\delta \to 0$ and hence $\delta w_\delta(y) \to \nu$ as $\delta \to 0$ uniformly on bounded sets of H. □

We notice that

$$\begin{aligned}\overline{F}(x, p, X) = \lim_{\delta\to 0} \sup_{a(\cdot)\in\mathcal{U}_0} \Bigg\{\mathbb{E}\int_0^{+\infty} \delta e^{-\delta s}\Bigg(&- \mathrm{Tr}(C(x, Y(s), a(s))X) \\ &- \langle b(x, Y(s), a(s)), p\rangle - l(x, Y(s), a(s))\Bigg)\mathrm{d}s\Bigg\},\end{aligned}$$

where $Y(s)$ is the solution of (5.1).

Remark 5.2. Differently from [1] and [4], since closed balls are on compact in H, we could not find a subsequence among $w_\delta - w_\delta(0)$ converging uniformly on bounded sets to pass to the limit in (5.2) in the proof of Theorem 5.1 and we had to resort to a comparison argument. However, if A_2 satisfies in addition $\langle A_2^* y, y\rangle \ge \lambda |y|_{2,1}^2$ for all $y \in D(A_2^*)$ then, using a definition of viscosity solution similar to Definition 4.7, one can show that the functions

$$\overline{w}(y) = \limsup_{\delta\to 0}\{w_\delta(y_\delta) - w_\delta(0) : y_\delta \to y\}, \quad \underline{w}(y) = \liminf_{\delta\to 0}\{w_\delta(y_\delta) - w_\delta(0) : y_\delta \to y\}$$

are respectively a (discontinuous) viscosity subsolution and a (discontinuous) viscosity supersolution of

$$\nu + \langle A_2 y, Dw\rangle + F\left(x, y, p, Dw, X, D^2 w, 0\right) = 0 \quad \text{in } H \tag{5.7}$$

and satisfy (5.3). Equation (5.7) then has a (discontinuous) viscosity solution by Perron's method (see [14, 19]). Moreover it can be proved that ν is the unique constant such that equation (5.7) has a (discontinuous) viscosity solution having at most linear growth.

6. Properties of the effective Hamiltonian $\overline{F}(x,p,X)$

We now prove properties of the effective Hamiltonian $\overline{F}(x,p,X)$. We rely on stochastic estimates.

Lemma 6.1. *Let Assumption 3.1 be satisfied. Let $x_1,x_2,y\in H$. Let $Y_1(s)$ be the solution of (5.1) with $x=x_1$ and $Y_2(s)$ be the solution of (5.1) with $x=x_2$. Then there exists a constant C, independent of $x_1,x_2\in H$ and $t\geq 0$, such that*

$$\mathbb{E}|Y_1(t)-Y_2(t)|^2\leq C|x_1-x_2|^2_{1,-1}\quad \textit{for all } t\geq 0. \tag{6.1}$$

Proof. Let $Y^1_N(s)$ be the solution of (5.5) with $x=x_1$ and initial condition $Y^1_N(0)=P_{2,N}y_1$ and $Y^2_N(s)$ be the solution of (5.5) with $x=x_1$ and initial condition $Y^2_N(0)=P_{2,N}y_2$. Applying Itô's formula and using (3.4) and (3.7) we get

$$\begin{aligned}
&e^{\mu t}\mathbb{E}|Y^1_N(t)-Y^2_N(t)|^2\leq \mu\mathbb{E}\int_0^t e^{\mu s}|Y^1_N(s)-Y^2_N(s)|^2\mathrm{d}s\\
&-2\mathbb{E}\int_0^t e^{\mu s}\langle A_2^*(Y^1_N(s)-Y^2_N(s)),Y^1_N(s)-Y^2_N(s)\rangle\mathrm{d}s\\
&+2\mathbb{E}\int_0^t e^{\mu s}\langle g(x_1,Y^1_N(s),a(s))-g(x_2,Y^1_N(s),a(s)),Y^1_N(s)-Y^2_N(s)\rangle\mathrm{d}s\\
&+2\mathbb{E}\int_0^t e^{\mu s}\langle g(x_2,Y^1_N(s),a(s))-g(x_2,Y^2_N(s),a(s)),Y^1_N(s)-Y^2_N(s)\rangle\mathrm{d}s\\
&+\mathbb{E}\int_0^t e^{\mu s}\mathrm{Tr}\big(((h(Y^1_N(s),a(s))-h(Y^2_N(s),a(s)))Q^{\frac12})\\
&\qquad\qquad ((h(Y^1_N(s),a(s))-h(Y^2_N(s),a(s)))Q^{\frac12})^*\big)\mathrm{d}s\\
&\leq \mu\mathbb{E}\int_0^t e^{\mu s}|Y^1_N(s)-Y^2_N(s)|^2\mathrm{d}s-2\mu\mathbb{E}\int_0^t e^{\mu s}|Y^1_N(s)-Y^2_N(s)|^2\mathrm{d}s\\
&+C\mathbb{E}\int_0^t e^{\mu s}|x_1-x_2|_{1,-1}|Y^1_N(s)-Y^2_N(s)|\mathrm{d}s\\
&\leq C\int_0^t e^{\mu s}|x_1-x_2|^2_{1,-1}\leq Ce^{\mu t}|x_1-x_2|^2_{1,-1}.
\end{aligned}$$

To estimate the trace term we used the fact that if $X,Y\in\mathcal{L}_2(K_0,H)$ then $XY^*\in\mathcal{L}_1(H)$ and $|\mathrm{Tr}(XY^*)|\leq \|XY^*\|_{\mathcal{L}_1(H)}\leq\|X\|_{\mathcal{L}_2(K_0,H)}\|Y\|_{\mathcal{L}_2(K_0,H)}$. □

Lemma 6.2. *Let Assumptions 3.1 be satisfied. The Hamiltonian $\overline{F}(x,p,X)$ has the following properties.*

– *For every $x,p\in H, X\in S(H)$,*

$$|\overline{F}(x,p,X)|\leq\sup_{y\in H}|F(x,y,p,0,X,0,0)|. \tag{6.2}$$

$$|\overline{F}(x,p,X)-\overline{F}(x,q,Y)|\leq C(|p-q|+\|X-Y\|)\quad\forall p,q\in H,X,Y\in S(H). \tag{6.3}$$

$$|\overline{F}(x_1,p,X)-\overline{F}(x_2,p,X)|\leq\omega(|x_1-x_2|)+C|x_1-x_2|(1+|p|+\|X\|)\quad\forall x_1,x_2,p\in H,X\in S(H). \tag{6.4}$$

– *There exists a modulus* ω_1 *such that*

$$\overline{F}\left(x_1, \frac{B_1(x_1-x_2)}{\varepsilon}, X\right) - \overline{F}\left(x_2, \frac{B_1(x_1-x_2)}{\varepsilon}, Y\right) \geq -\omega_1\left(|x_1-x_2|\left(1+\frac{|x_1-x_2|_{1,-1}}{\varepsilon}\right)\right) \tag{6.5}$$

for every $\varepsilon > 0, x_1, x_2 \in H, X, Y \in S(H), X = P_{1,N}XP_{1,N}, Y = P_{1,N}YP_{1,N}$,

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\epsilon}\begin{pmatrix} B_1P_{1,N} & -B_1P_{1,N} \\ -B_1P_{1,N} & B_1P_{1,N} \end{pmatrix}. \tag{6.6}$$

– *There exists a modulus* ω_2 *such that*

$$\overline{F}\left(x_1, \frac{x_1-x_2}{\varepsilon}, X\right) - \overline{F}\left(x_2, \frac{x_1-x_2}{\varepsilon}, Y\right) \geq -\omega_2\left(|x_1-x_2|\left(1+\frac{|x_1-x_2|}{\varepsilon}\right)\right) \tag{6.7}$$

for every $\varepsilon > 0, x_1, x_2 \in H, X, Y \in S(H), X = P_{1,N}XP_{1,N}, Y = P_{1,N}YP_{1,N}$,

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\epsilon}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix}. \tag{6.8}$$

– *For all* $x, p \in H, R > 0$,

$$\sup\left\{|\overline{F}(x,p,X+\kappa Q_{1,N}) - \overline{F}(x,p,X)| : |\kappa| \leq R, X = P_{1,N}XP_{1,N}\right\} \to 0 \ \ as \ \ N\to\infty. \tag{6.9}$$

Proof. Inequality (6.2) is obtained arguing as in the proof of Theorem 5.1, since we have $|\delta w_\delta| \leq M$, where M is defined by (5.4). This gives (6.2) by the definition of $\overline{F}(x,p,X)$.

Inequality (6.3) is obvious from the definition of $\overline{F}(x,p,X)$. Estimate (6.4) is a simple consequence of (3.5), (3.6), (3.7), (3.8) and (6.1). Inequality (6.5) also follows from (3.5), (3.6), (3.7), (3.8) and (6.1). Using the notation of Lemma 6.1, for every control $a(\cdot)$ we have

$$\begin{aligned}&\mathbb{E}\left|\langle b(x_1,Y_1(s),a(s)), \frac{B_1(x_1-x_2)}{\varepsilon}\rangle - \langle b(x_2,Y_2(s),a(s)), \frac{B_1(x_1-x_2)}{\varepsilon}\rangle\right| \\ &\leq \mathbb{E}L(|x_1-x_2| + |Y_1(t)-Y_2(t)|)\frac{|B_1(x_1-x_2)|}{\varepsilon} \leq C|x_1-x_2|\frac{|x_1-x_2|_{1,-1}}{\varepsilon}.\end{aligned}$$

Similar argument gives

$$\mathbb{E}|l(x_1,Y_1(s),a(s)) - l(x_2,Y_2(s),a(s))| \leq C|x_1-x_2|_{1,-1} + \omega(|x_1-x_2|).$$

Moreover, arguing as in [14], pages 235–236, we obtain

$$\begin{aligned}
&\mathrm{Tr}(C(x_1,Y_1(s),a(s))X)-\mathrm{Tr}(C(x_2,Y_2(s),a(s))Y)\\
&\le \frac{3}{\varepsilon}\mathrm{Tr}\big(((\sigma(x_1,Y_1(s),a(s))-\sigma(x_2,Y_2(s),a(s)))Q^{\frac12})((\sigma(x_1,Y_1(s),a(s))-\sigma(x_2,Y_2(s),a(s)))Q^{\frac12})^*B_1\big)\\
&\le C(|x_1-x_2|_{1,-1}+|Y_1(t)-Y_2(t)|)^2.
\end{aligned}$$

This, together with (6.1), gives the required estimate for the trace terms. The three inequalities above give (6.5). Inequality (6.7) is proved similarly. The convergence in (6.9) is a consequence of (3.9) and the definition of $\overline{F}$. □

7. Convergence of viscosity solutions

In this section we prove the main result, Theorem 7.2, about convergence of the viscosity solutions u_ε of (1.2) to the viscosity solution of the effective equation (7.1). We need first to prove the existence of a unique viscosity solution of the limiting effective equation (7.1) and show comparison principle for discontinuous viscosity solutions of (7.1).

Theorem 7.1. *Let $\overline{F}$ satisfy the properties stated in Lemma 6.2 and f satisfy (3.7). Then there exists a unique bounded viscosity solution $u\in UB_b^x((0,T]\times H)\cap UB_b^x((0,\tau]\times H_{1,-1})$ for $0<\tau<T$, in the sense of Definition 4.4 of the equation*

$$\begin{cases}\partial_t u-\langle A_1x,Du\rangle-\overline{F}\left(x,Du,D^2u\right)=0 & \text{in } (0,T)\times H,\\ u(T,x)=f(x) & \text{for } x\in H\end{cases}\tag{7.1}$$

satisfying

$$\lim_{t\to T,t<T}\sup_{x\in H}|(u(t,x)-f(e^{-(T-t)A_1}x))|=0.\tag{7.2}$$

Moreover if u is a bounded discontinuous viscosity subsolution of (7.1) and v is a bounded discontinuous viscosity supersolution of (7.1) satisfying

$$\lim_{t\to T,t<T}\left[(u(t,x)-f(e^{-(T-t)A_1}x))^+ +(v(t,x)-f(e^{-(T-t)A_1}x))^-\right]=0\tag{7.3}$$

uniformly on bounded subsets of H, then $u\le v$ on $(0,T]\times H$.

Proof. The existence of a unique viscosity solution u of (7.1) in the sense of Definition 3.35 of [14] with the required properties follows from Theorem 3.84 of [14]. Uniqueness can also be deduced from arguments in [14], however it will follow from an even stronger comparison principle for discontinuous viscosity solutions which we will prove below. To show that u is also a viscosity solution of (7.1) in the sense of Definition 4.4 we recall that the solution in Theorem 3.84 of [14] was obtained as the limit as $N\to\infty$, uniform on bounded subsets of $(0,\tau)\times H$ for every $0<\tau<T$, of the viscosity solutions $u_N, N\ge 1$, of the approximating problems

$$\begin{cases}\partial_t u_N-\langle A_{1,N}x,Du_N\rangle-\overline{F}\left(P_{1,N}x,P_{1,N}Du_N,P_{1,N}D^2u_NP_{1,N}\right)=0 & \text{in } (0,T)\times H,\\ u_N(T,x)=f(P_{1,N}x) & \text{for } x\in H,\end{cases}\tag{7.4}$$

where $A_{1,N}=(P_{1,N}A_1^*P_{1,N})^*$. We have $\langle A_{1,N}x,x\rangle\ge\lambda|P_{1,N}x|_{1,1}^2$.

Suppose now that $u-\psi$ has a strict global maximum at (t,x) for a test function $\psi(s,y)=\eta(s)\theta(|y|)+\varphi(s,y)$. Then there is a sequence $(t_N,x_N)\to(t,x)$ such that $u_N-\psi$ has a local maximum at (t_N,x_N). Therefore we have

$$\begin{aligned}0 &\le \partial_t\psi(t_N,x_N)-\eta(t_N)\frac{\theta'(|x_N|)}{|x_N|}\langle A_{1,N}x_N,x_N\rangle-\langle A_{1,N}x_N,D\varphi(t_N,x_N)\rangle\\ &\quad-\overline{F}\left(P_{1,N}x_N,P_{1,N}D\psi(t_N,x_N),P_{1,N}D^2\psi(t_N,x_N)P_{1,N}\right)\\ &\le \partial_t\psi(t_N,x_N)-\lambda\eta(t_N)\frac{\theta'(|x_N|)}{|x_N|}|P_{1,N}x_N|_{1,1}^2-\langle P_{1,N}x_N,A_1^*P_{1,N}D\varphi(t_N,x_N)\rangle\\ &\quad-\overline{F}\left(P_{1,N}x_N,P_{1,N}D\psi(t_N,x_N),P_{1,N}D^2\psi(t_N,x_N)P_{1,N}\right).\end{aligned}\tag{7.5}$$

It follows from Lemma 3.83(i) of [14] that $A_1^*P_{1,N}D\varphi(t_N,x_N)\to A_1D\varphi(t,x)$. Therefore, since $\overline{F}$ is continuous, $\eta>0$, $\theta'>0$ and $\theta''(0)>0$ (in case $x=0$), (7.5) implies

$$|P_{1,N}x_N|_{1,1}^2\le C\quad\text{for all } N\ge 1.$$

Thus $P_{1,N}x_N\rightharpoonup\bar{x}$ in $H_{1,1}$ for some $\bar{x}\in H_{1,1}$. But $P_{1,N}x_N\to x$ in H so we must have $x=\bar{x}\in H_{1,1}$. We can now pass to the limit as $N\to\infty$ using the continuity properties of $\overline{F}$ and Lemma 3.85 of [14] to get

$$\partial_t\psi(t,x)-\lambda\eta(t)\frac{\theta'(|x)}{|x|}|x|_{1,1}^2-\langle x,A_1^*D\varphi(t,x)\rangle-\overline{F}\left(x,D\psi(t,x),D^2\psi(t,x)\right)\ge 0.$$

The proof of the supersolution property is similar.

We now prove the comparison statement for bounded discontinuous viscosity subsolutions and supersolutions. We argue by contradiction and assume that $u\not\le v$. Then there is $\gamma>0$ such that for every sufficiently small $\mu>0$ and every $T_1<T$ sufficiently close to T,

$$\begin{aligned}\gamma<m(T_1)&:=\lim_{R\uparrow\infty}\lim_{(r,\alpha)\downarrow(0,0)}\sup\{u^\mu(t,x)-v^\mu(s,y):\\ &\qquad |t-s|<\alpha,|x-y|_{1,-1}<r,x,y\in B(0,R),0<t,s\le T_1\}\\ &=\lim_{R\uparrow\infty}\lim_{r\downarrow 0}\lim_{\alpha\downarrow 0}\sup\{u^\mu(t,x)-v^\mu(s,y):\\ &\qquad |t-s|<\alpha,|x-y|_{1,-1}<r,x,y\in B(0,R),0<t,s\le T_1\},\end{aligned}\tag{7.6}$$

where we have set $u^\mu(t,x)=u(t,x)-\frac{\mu}{t}$ and $v^\sigma(s,y)=v(s,y)+\frac{\mu}{s}$. We define $u_\eta(t,x)=u^\mu(t,x)-\eta e^{K(T-t)}|x|^2$, $v_\eta(s,y)=v^\mu(s,y)+\eta e^{K(T-s)}|y|^2$ and $K=C$, where C is from (6.3). Let

$$\begin{aligned}m_\eta:=&\lim_{r\downarrow 0}\lim_{\alpha\downarrow 0}\sup\{(u_\eta)^*(t,x)-(v_\eta)_*(s,y):\,|x-y|_{1,-1}<r,|t-s|<\alpha,0<t,s\le T_1\}\\ =&\lim_{r\downarrow 0}\lim_{\alpha\downarrow 0}\sup\{u_\eta(t,x)-v_\eta(s,y):\,|x-y|_{1,-1}<r,|t-s|<\alpha,0<t,s\le T_1\},\end{aligned}$$

where $(u_\eta)^*$ and $(v_\eta)_*$ are computed using only points in $(0,T_1]\times H$,

$$m_{\eta,\epsilon}:=\lim_{\alpha\downarrow 0}\sup\left\{(u_\eta)^*(t,x)-(v_\eta)_*(s,y)-\frac{1}{2\varepsilon}|x-y|_{1,-1}^2:\,|t-s|<\alpha,0<t,s\le T_1\right\},$$

$$m_{\eta,\varepsilon,\beta} := \sup\left\{(u_\eta)^*(t,x) - (v_\eta)_*(s,y) - \frac{1}{2\varepsilon}|x-y|^2_{1,-1} - \frac{(t-s)^2}{2\beta} \,:\, 0 < t,s \le T_1\right\}.$$

We have

$$m(T_1) = \lim_{\eta\downarrow 0} m_\eta, \tag{7.7}$$

$$m_\eta = \lim_{\varepsilon\downarrow 0} m_{\eta,\varepsilon}, \tag{7.8}$$

$$m_{\eta,\epsilon} = \lim_{\beta\downarrow 0} m_{\eta,\varepsilon,\beta}, \tag{7.9}$$

Since B_1 is compact, the function

$$(u_\eta)^*(t,x) - (v_\eta)_*(s,y) - \frac{1}{2\varepsilon}|x-y|^2_{1,-1} - \frac{1}{2\beta}|t-s|^2$$

is weakly sequentially upper semicontinuous on $[0,T_1]\times[0,T_1]\times H\times H$. Thus it achieves a global maximum at some point $(\bar t,\bar s,\bar x,\bar y)\in[0,T_1]\times[0,T_1]\times H\times H$ which can be assumed to be strict in the $|\cdot|\times|\cdot|\times|\cdot|_{1,-1}\times|\cdot|_{1,-1}$ norm. By Definition 4.7, $\bar x,\bar y\in H_{1,1}$. Convergences (7.7)–(7.9) yield (see for instance [14], page 208 for such arguments)

$$\lim_{\beta\downarrow 0}\frac{1}{2\beta}|\bar t-\bar s|^2 = 0 \quad \text{for every } \eta,\varepsilon>0, \tag{7.10}$$

$$\lim_{\varepsilon\downarrow 0}\limsup_{\beta\downarrow 0}\frac{1}{2\varepsilon}|\bar x-\bar y|^2_{1,-1} = 0 \quad \text{for every } \eta>0. \tag{7.11}$$

It follows from the definitions of u_η and v_η that $\bar t,\bar s>0$ and that $|\bar x|,|\bar y|\le R_\eta$ for some $R_\eta>0$ independent of ε,β,μ and T_1. Suppose that we also have $\bar t,\bar s<T_1$. We will show that this is impossible. We set

$$u_1(t,x) = (u_\eta)^*(t,x) - \frac{\langle B_1 Q_{1,N}(\bar x-\bar y),x\rangle}{\varepsilon} - \frac{|Q_{1,N}(x-\bar x)|^2_{1,-1}}{\varepsilon} + \frac{|Q_{1,N}(\bar x-\bar y)|^2_{-1,1}}{2\varepsilon},$$

$$v_1(s,y) = (v_\eta)_*(s,y) - \frac{\langle BQ_{1,N}(\bar x-\bar y),y\rangle}{\varepsilon} + \frac{|Q_{1,N}(y-\bar y)|^2_{1,-1}}{\varepsilon}.$$

Then

$$u_1(t,x) - v_1(s,y) - \frac{1}{2\varepsilon}|P_{1,N}(x-y)|^2_{1,-1} - \frac{1}{2\beta}|t-s|^2$$

has a strict (in the $|\cdot|\times|\cdot|\times|\cdot|_{1,-1}\times|\cdot|_{1,-1}$ norm) global maximum at $(\bar t,\bar s,\bar x,\bar y)$. We can now use Theorem 3.27 and Corollary 3.29 of [14]. However our maximum is not strict in the standard norm. This will result in one change in the statement of Corollary 3.29 which follows trivially from the proof of Theorem 3.27 of [14]. (In fact the proofs of Theorem 3.27 and Corollary 3.29 of [14] are easier in our case since all the functions involved are weakly sequentially upper semicontinuous and so there is no need to use perturbed optimization). We leave the details to the readers. We thus obtain that there exist test functions φ_k,ψ_k satisfying the properties of Definition 4.3-(ii) such that

$$u_1(t,x) - \varphi_k(t,x)$$

has a global maximum at some point (t_k, x_k),

$$v_1(s,y) + \psi_k(s,y)$$

has a global minimum at some point (s_k, y_k), and

$$\begin{aligned}&\big(t_k, x_k, u_1(t_k,x_k), \partial_t\varphi_k(t_k,x_k), D\varphi_k(t_k,x_k), D^2\varphi_k(t_k,x_k)\big)\\ &\qquad \to \big(\bar{t}, \bar{x}, u_1(\bar{t},\bar{x}), \frac{\bar{t}-\bar{s}}{\beta}, \frac{BP_{1,N}(\bar{x}-\bar{y})}{\varepsilon}, X_N\big),\end{aligned} \tag{7.12}$$

$$\begin{aligned}&\big(s_k, y_k, v_1(s_k,y_k), \partial_t\psi_k(s_k,y_k), D\psi_k(s_k,y_k), D^2\psi_k(s_k,y_k)\big)\\ &\qquad \to \big(\bar{s}, \bar{y}, v_1(\bar{s},\bar{y}), \frac{\bar{s}-\bar{t}}{\beta}, \frac{BP_{1,N}(\bar{y}-\bar{x})}{\varepsilon}, Y_N\big),\end{aligned} \tag{7.13}$$

$$-\frac{3}{\varepsilon}\begin{pmatrix} B_1P_{1,N} & 0 \\ 0 & B_1P_{1,N}\end{pmatrix} \le \begin{pmatrix} X_N & 0 \\ 0 & Y_N\end{pmatrix} \le \frac{3}{\varepsilon}\begin{pmatrix} B_1P_{1,N} & -B_1P_{1,N} \\ -B_1P_{1,N} & B_1P_{1,N}\end{pmatrix}, \tag{7.14}$$

with the convergences being in $\mathbb{R}\times H_{1,-1}\times\mathbb{R}\times\mathbb{R}\times H_{1,2}\times\mathcal{L}(H)$ and moreover the sequences of points $x_k, y_k, k = 1,2,...$ are bounded.

We have from the definition of a viscosity subsolution and supersolution that $x_k, y_k \in H_{1,1}$. Moreover

$$\begin{aligned}&-\frac{\mu}{t_k} + \partial_t\varphi(t_k,x_k) - \eta K e^{K(T-t_k)}|x_k|^2\\ &-2\lambda\eta|x_k|^2_{1,1} - \left\langle x_k, A_1^*\left(D\varphi_k(t_k,x_k) + \frac{B_1Q_{1,N}(\bar{x}-\bar{y})}{\varepsilon} + \frac{2B_1Q_{2,N}(x_k-\bar{x})}{\varepsilon}\right)\right\rangle\\ &-\overline{F}\left(x_k, 2\eta e^{K(T-t_k)}x_k + D\varphi_k(t_k,x_k) + \frac{B_1Q_{1,N}(\bar{x}-\bar{y})}{\varepsilon} + \frac{2B_1Q_{2,N}(x_k-\bar{x})}{\varepsilon},\right.\\ &\qquad\qquad \left. 2\eta e^{K(T-t_k)}I + D^2\varphi_k(t_k,x_k) + \frac{2B_1Q_{2,N}}{\varepsilon}\right) \ge 0.\end{aligned} \tag{7.15}$$

We deduce from (7.15) that $|B_1^{-\frac{1}{2}}x_k|^2 = |x_k|^2_{1,1} \le M$ for some M and all $k \ge 1$. Since $B_1^{\frac{1}{2}}x_k \to B_1^{\frac{1}{2}}\bar{x}$, we have

$$|x_k-\bar{x}|^2 = \langle B_1^{\frac{1}{2}}(x_k-\bar{x}), B_1^{-\frac{1}{2}}(x_k-\bar{x})\rangle \to 0 \quad \text{as } k\to\infty.$$

Using this, (7.12), (7.13) and dropping the $2\lambda\eta|x_k|^2_{1,1}$ term, we pass to the limit as $k\to\infty$ in (7.15) to obtain

$$\begin{aligned}&-\frac{\mu}{\bar{t}} - \eta K e^{K(T-\bar{t})}|\bar{x}|^2 + \frac{\bar{t}-\bar{s}}{\beta} - \left\langle \bar{x}, \frac{A_1^*B_1(\bar{x}-\bar{y})}{\varepsilon}\right\rangle\\ &-\overline{F}\left(\bar{x}, 2\eta e^{K(T-\bar{t})}\bar{x} + \frac{B_1(\bar{x}-\bar{y})}{\varepsilon}, 2\eta e^{K(T-\bar{t})}I + X_N + \frac{2B_1Q_{2,N}}{\varepsilon}\right) \ge 0.\end{aligned} \tag{7.16}$$

Using (6.3) we now have

$$-\frac{\mu}{T}+\frac{\bar t-\bar s}{\beta}-\left\langle \bar x,\frac{A_1^*B_1(\bar x-\bar y)}{\varepsilon}\right\rangle-\overline{F}\left(\bar x,\frac{B_1(\bar x-\bar y)}{\varepsilon},X_N+\frac{2B_1Q_{2,N}}{\varepsilon}\right)\geq -3\eta Ke^{K(T-\bar t)}$$

which, upon employing (6.9), yields

$$-\frac{\mu}{T}+\frac{\bar t-\bar s}{\beta}-\left\langle \bar x,\frac{A_1^*B_1(\bar x-\bar y)}{\varepsilon}\right\rangle-\overline{F}\left(\bar x,\frac{B_1(\bar x-\bar y)}{\varepsilon},X_N\right)\geq -3\eta Ke^{KT}-\omega_\varepsilon(N) \tag{7.17}$$

where, for ε fixed, $\omega_\varepsilon(N)\to 0$ as $N\to\infty$.

Similar arguments produce

$$\frac{\mu}{T}+\frac{\bar t-\bar s}{\beta}-\left\langle \bar y,\frac{A_1^*B_1(\bar x-\bar y)}{\varepsilon}\right\rangle-\overline{F}\left(\bar y,\frac{B_1(\bar x-\bar y)}{\varepsilon},-Y_N\right)\leq 3\eta Ke^{KT}+\omega_\varepsilon(N). \tag{7.18}$$

Subtracting (7.17) from (7.18) we thus obtain

$$\begin{aligned}\frac{2\mu}{T}+\left\langle \bar x-\bar y,\frac{A_1^*B_1(\bar x-\bar y)}{\varepsilon}\right\rangle+\overline{F}\left(\bar x,\frac{B_1(\bar x-\bar y)}{\varepsilon},X_N\right)-\overline{F}\left(\bar y,\frac{B_1(\bar x-\bar y)}{\varepsilon},-Y_N\right)&\\ \leq 6\eta Ke^{K(T-\bar t)}+2\omega_\varepsilon(N).&\end{aligned}$$

In now follows from (3.3) and (6.5) that

$$\begin{aligned}\frac{2\mu}{T}&\leq 6\eta Ke^{KT}+2\omega_\varepsilon(N)\\ &\quad+\omega_1\left(|\bar x-\bar y|\left(1+\frac{|\bar x-\bar y|_{1,-1}}{\varepsilon}\right)\right)+\frac{c}{\varepsilon}|\bar x-\bar y|_{1,-1}^2-\frac{1}{\varepsilon}|\bar x-\bar y|^2.\end{aligned} \tag{7.19}$$

If $\omega_1(s)\leq\frac{\mu}{T}+M_\mu s, s>0$, an elementary computation shows that for some constant C_μ

$$\omega_1\left(|\bar x-\bar y|\left(1+\frac{|\bar x-\bar y|_{1,-1}}{\varepsilon}\right)\right)\leq\frac{\mu}{T}+C_\mu\varepsilon+\frac{C_\mu}{\varepsilon}|\bar x-\bar y|_{1,-1}^2+\frac{1}{\varepsilon}|\bar x-\bar y|^2. \tag{7.20}$$

Plugging (7.20) into (7.19) now gives us

$$\frac{\mu}{T}\leq 6\eta Ke^{KT}+2\omega_\varepsilon(N)+C_\mu\varepsilon+\frac{C_\mu}{\varepsilon}|\bar x-\bar y|_{1,-1}^2. \tag{7.21}$$

It remains to take $\lim_{\eta\to 0}\lim_{\varepsilon\to 0}\limsup_{\beta\to 0}\lim_{N\to\infty}$ in (7.21) and use (7.11) to obtain a contradiction.

Therefore we must have that either $\bar t=T_1$ or $\bar s=T_1$. It follows from (7.6)–(7.9) that for every T_1 sufficiently close to T and sufficiently small $\mu,\eta,\varepsilon,\beta$,

$$(u_\eta)^*(\bar t,\bar x)-(v_\eta)_*(\bar s,\bar y)>\gamma>0.$$

Thus, by the definition of $(u_\eta)^*$ and $(v_\eta)_*$ and (7.10), (7.11), there exist points $(\tilde t,\tilde s,\tilde x,\tilde y)\in(0,T_1]\times(0,T_1]\times H\times H$, depending on $\mu,\eta,\varepsilon,\beta$, such that $T_1-\tilde s\leq C\sqrt{\beta},T_1-\tilde t\leq C\sqrt{\beta}$, $|\tilde x-\tilde y|_{1,-1}\leq C\sqrt{\varepsilon}$, $|\tilde x|,|\tilde y|\leq R_\eta$ for some $R_\eta>0$ independent of $\mu,\eta,\varepsilon,\beta,T_1$, and such that

$$u(\tilde t,\tilde x)-v(\tilde s,\tilde y)>\gamma.$$

Denote $\tau = T - T_1$. Without loss of generality we will suppose that $\tilde{s} \geq \tilde{t}$. Then, by (7.3),

$$\begin{aligned}\gamma &< \left(u(\tilde{t},\tilde{x}) - f(e^{-(T-\tilde{t})A_1}\tilde{x})\right)^+ + \left(f(e^{-(T-\tilde{t})A_1}\tilde{x}) + f(e^{-(T-\tilde{t})A_1}\tilde{y})\right)\\ &+ \left(f(e^{-(T-\tilde{t})A_1}\tilde{y}) + f(e^{-(T-\tilde{s})A_1}\tilde{y})\right) + \left(v(\tilde{s},\tilde{y}) - f(e^{-(T-\tilde{s})A_1}\tilde{y})\right)^-\\ &\leq \omega_\eta\left(\tau + C\sqrt{\beta}\right) + \omega\left(|e^{-(T-\tilde{t})A_1}(\tilde{x}-\tilde{y})|\right) + \omega\left(|e^{-(T-\tilde{s})A_1}(e^{-(\tilde{s}-\tilde{t})A_1}\tilde{y}-\tilde{y})|\right),\end{aligned} \tag{7.22}$$

where ω_η are some moduli coming from (7.3) and ω is the modulus from (3.7). By Lemma 3.19

$$\omega\left(|e^{-(T-\tilde{t})A_1}(\tilde{x}-\tilde{y})|\right) \leq \omega\left(\frac{C_1}{\sqrt{\tau}}|\tilde{x}-\tilde{y}|_{1,-1}\right) \leq \omega\left(\frac{C_2\sqrt{\varepsilon}}{\sqrt{\tau}}\right)$$

and

$$\omega\left(|e^{-(T-\tilde{s})A_1}(e^{-(\tilde{s}-\tilde{t})A_1}\tilde{y}-\tilde{y})|\right) \leq \omega\left(\frac{C_1}{\sqrt{\tau}}|e^{-(\tilde{s}-\tilde{t})A_1}\tilde{y}-\tilde{y}|_{1,-1}\right) \leq \omega\left(\frac{C_\eta\sqrt{\beta}}{\sqrt{\tau}}\right)$$

Plugging these into (7.22) and taking $\lim_{\varepsilon\to 0}\lim_{\beta\to 0}$ we thus obtain

$$\gamma \leq \omega_\eta(\tau)$$

which is again a contradiction for sufficiently small τ. Therefore we must have $u \leq v$. □

We can now prove the main result of the paper. The overall idea of the proof of Theorem 7.2 is similar to that of the proof of Theorem 14 of [1] in the finite dimensional case.

Theorem 7.2. *Let Assumption 3.1 be satisfied. Then, for every $0 < T_1 < T, R > 0$ and every compact set $K \subset H$, the family $\{u_\varepsilon\}$ converges uniformly on $[T_1, T] \times K \times B(0,R)$, as $\varepsilon \to 0$, to the unique bounded viscosity solution u of (7.1) satisfying (7.2).*

Proof. It follows from comparison that if

$$M_1 = \sup_{x\in H}|f(x)|, \quad M_2 = \sup_{x,y\in H}|F(x,y,0,0,0,0,0)|$$

then for every $\varepsilon > 0$,

$$|u_\varepsilon| \leq M_1 + M_2(T-t).$$

Define

$$F_1(x,p,X) = \sup_{y\in H} F(x,y,p,0,X,0,0), \quad F_2(x,p,X) = \inf_{y\in H} F(x,y,p,0,X,0,0).$$

The functions F_1, F_2 satisfy the same assumptions as $\overline{F}$ so (by Thm. 7.1), for $i = 1, 2$, the equations

$$\begin{cases} \partial_t w_i - \langle A_1 x, DW_i\rangle - F_i\left(x, DW_i, D^2 w_i\right) = 0 & \text{in } (0,T)\times H,\\ w_i(T,x) = f(x) & \text{for } x \in H\end{cases} \tag{7.23}$$

have unique bounded viscosity solutions $w_i \in UB_b^x((0,T]\times H)\cap UB_b^x((0,\tau]\times H_{1,-1})$ for $0<\tau<T$, satisfying

$$\lim_{t\to T, t<T}\sup_{x\in H}|(w_i(t,x)-f(e^{-(T-t)A_1}x))|=0. \tag{7.24}$$

It is easy to see that w_1 is a viscosity supersolution of (1.2) and w_2 is a viscosity subsolution of (1.2). Therefore, by comparison, we get $w_2\le u_\epsilon\le w_1$ on $(0,T]\times H\times H$ for every $\varepsilon>0$.

We define similarly as in [15]

$$\underline{u}(t,x)=\inf\liminf_{\varepsilon\to 0}\left\{u_\varepsilon(t_\varepsilon,x_\varepsilon,y_\varepsilon): t_\varepsilon\to t, x_\varepsilon\to x, y_\varepsilon \text{ is bounded}\right\},$$

$$\overline{u}(t,x)=\sup\limsup_{\varepsilon\to 0}\left\{u_\varepsilon(t_\varepsilon,x_\varepsilon,y_\varepsilon): t_\varepsilon\to t, x_\varepsilon\to x, y_\varepsilon \text{ is bounded}\right\}.$$

Since $w_2\le u_\epsilon\le w_1$ on $(0,T]\times H\times H$, we have $w_2\le\underline{u}\le\overline{u}\le w_1$ on $(0,T]\times H$, so in particular $\overline{u}(T,x)=\underline{u}(T,x)=f(x)$ and the functions $\underline{u},\overline{u}$ satisfy (7.24). We claim that $\underline{u}$ is a discontinuous viscosity supersolution of (7.1) and $\overline{u}$ is a discontinuous viscosity subsolution of (7.1). We will only prove the subsolution statement as the supersolution claim is proved similarly.

Denote $\overline{F}=\overline{F}(\bar{x},D\psi(\bar{t},\bar{x}),D^2\psi(\bar{t},\bar{x}))$. Let v_ε be the viscosity solution of

$$\varepsilon v_\varepsilon+\langle A_2y,D_yv_\varepsilon\rangle+F\left(\bar{x},y,D\psi(\bar{t},\bar{x}),D_yv_\varepsilon,D^2\psi(\bar{t},\bar{x}),D_y^2v_\varepsilon,0\right)-\overline{F}=0\quad\text{in } H.$$

We know by Theorem 5.1 and the definition of $\overline{F}$, that $\varepsilon(v_\varepsilon-\varepsilon^{-1}\overline{F})$ converges uniformly on bounded sets of H to $-\overline{F}$. This implies that $\varepsilon v_\varepsilon$ converges uniformly on bounded sets of H to 0. It also follows from comparison that $|\varepsilon v_\varepsilon|\le K$ for some K and all ε.

Suppose that $(\overline{u}-\eta(\cdot)\theta(|\cdot|))^*-\varphi$ has a strict in $|\cdot|\times|\cdot|_{1,-1}$ norm maximum over $(0,T]\times H$ at some point $(\bar{t},\bar{x})\in(0,T)\times H$. We can assume that $(\overline{u}-\eta(\cdot)\theta(|\cdot|))^*(\bar{t},\bar{x})-\varphi(\bar{t},\bar{x})=0$. By modifying the functions θ and φ if necessary we can assume that $(\overline{u}-\eta(\cdot)\theta(|\cdot|))^*(t,x)-\varphi(t,x)\to-\infty$ as $|x|\to\infty$, uniformly for $t\in(0,T]$, for every ε, $u_\varepsilon(t,x,y)-\eta(t)\theta(|x|)-\varphi(t,x)\to-\infty$ as $|x|\to\infty$, uniformly for $(t,y)\in(0,T]\times H$, and there exist $R>|\bar{x}|, r>0$, such that if $u_\varepsilon(t,x,y)-\eta(t)\theta(|x|)-\varphi(t,x)>-(K+1)$, then $(t,x)\in(\bar{t}-r,\bar{t}+r)\times\big(B_{H_{-1}}(\bar{x},r)\cap B(0,R)\big)$.

It now follows that there must exist a sequence $(\tilde{t}_n,\tilde{x}_n)$ such that $\tilde{t}_n\to\bar{t}, |\tilde{x}_n-\bar{x}|_{1,-1}\to 0, |\tilde{x}_n|\le C$ and

$$\overline{u}(\tilde{t}_n,\tilde{x}_n)-\eta(\tilde{t}_n)\theta(|\tilde{x}_n|)-\varphi(\tilde{t}_n,\tilde{x}_n)\ge-\frac{1}{n}.$$

Therefore, for every $\delta>0$, there exist $\varepsilon_n\to 0$ and $(\tilde{s}_n,\tilde{z}_n,\tilde{y}_n)$ such that

$$u_{\varepsilon_n}(\tilde{s}_n,\tilde{z}_n,\tilde{y}_n)-\varepsilon_n v_{\varepsilon_n}(\tilde{y}_n)-\delta\varepsilon_n|\tilde{y}_n|^2-\eta(\tilde{s}_n)\theta(|\tilde{z}_n|)-\varphi(\tilde{s}_n,\tilde{z}_n)\ge-\frac{2}{n}.$$

Let (t_n,x_n,y_n) be a global maximum over $(0,T]\times H\times H$ of the function $\Psi(t,x,y)=u_{\varepsilon_n}(t,x,y)-\varepsilon_n v_{\varepsilon_n}(y)-\delta\varepsilon_n|y|^2-\eta(t)\theta(|x|)-\varphi(t,x)$ which we can assume to be strict. We must have $(t_n,x_n)\in(\bar{t}-r,\bar{t}+r)\times\big(B_{H_{-1}}(\bar{x},r)\cap B(0,R)\big)$. We want to produce appropriate test functions for the functions u_{ε_n} and v_{ε_n}. Recall that the test functions must have a nontrivial radial component. Thus we will introduce an extra variable z and a parameter $\alpha>0$, split the function $\delta\varepsilon_n|y|^2$ into two parts for the variables y and z, and then penalize the doubling by subtracting the term $\frac{|z-y|_{2,-1}^2}{2\alpha}$.

We thus fix n and, for $\alpha > 0$, let $(t_\alpha, x_\alpha, z_\alpha, y_\alpha)$ be a global maximum over $(0,T] \times H \times H \times H$ of

$$\Psi_\alpha(t,x,z,y) = u_{\varepsilon_n}(t,x,z) - \frac{\delta\varepsilon_n}{2}|z|^2 - \eta(t)\theta(|x|) - \varphi(t,x) - \varepsilon_n v_{\varepsilon_n}(y) - \frac{\delta\varepsilon_n}{2}|y|^2 - \frac{|z-y|^2_{2,-1}}{2\alpha}$$

which we can also consider to be strict. We know that $x_\alpha \in H_{1,1}$. It is easy to see that we must have $|x_\alpha|, |z_\alpha|, |y_\alpha| \le C$,

$$\lim_{\alpha\to 0} \Psi_\alpha(t_\alpha, x_\alpha, z_\alpha, y_\alpha) = \Psi(t_n, x_n, y_n)$$

and

$$\lim_{\alpha\to 0} \frac{|z_\alpha - y_\alpha|^2_{2,-1}}{2\alpha} = 0. \tag{7.25}$$

Since $u_{\varepsilon_n}, v_{\varepsilon_n}, \varphi$ are weakly sequentially continuous, the norm is weakly sequentially lower semicontinuous, and Ψ has a strict maximum at (t_n, x_n, y_n), passing to a subsequence we can assume that $t_\alpha \to t_n, x_\alpha \rightharpoonup x_n, z_\alpha \rightharpoonup y_n, y_\alpha \rightharpoonup y_n$ but then we must also have $|x_\alpha| \to |x_n|, |z_\alpha| \to |y_n|, |y_\alpha| \to |y_n|$ so we conclude that $(t_\alpha, x_\alpha, z_\alpha, y_\alpha) \to (t_n, x_n, y_n, y_n)$ as $\alpha \to 0$. In particular, we can assume that $(t_\alpha, x_\alpha) \in (\bar t - r, \bar t + r) \times \big(B_{H_{-1}}(\bar x, r) \cap B(0,R)\big)$.

We now have

$$|z-y|^2_{2,-1} = |P_{2,N}(z-y)|^2_{2,-1} + |Q_{2,N}(z-y)|^2_{2,-1}$$

and

$$\begin{aligned}|Q_{2,N}(z-y)|^2_{2,-1} \le 2\langle B_2 Q_{2,N}(z_\alpha - y_\alpha), z-y\rangle + 2|Q_{2,N}(z-z_\alpha)|^2_{2,-1} \\ + 2|Q_{2,N}(y-y_\alpha)|^2_{2,-1} - |Q_{2,N}(z_\alpha - y_\alpha)|^2_{2,-1}\end{aligned}$$

with equality at z_α, y_α. Thus, defining

$$\begin{aligned}U_n(t,x,z) = u_{\varepsilon_n}(t,x,z) - \eta(t)\theta(|x|) - \varphi(t,x) - \frac{\delta\varepsilon_n}{2}|z|^2 \\ - \frac{\langle B_2 Q_{2,N}(z_\alpha - y_\alpha), z\rangle}{\alpha} - \frac{|Q_{2,N}(z-z_\alpha)|^2_{2,-1}}{\alpha} + \frac{|Q_{2,N}(z_\alpha - y_\alpha)|^2_{2,-1}}{2\alpha}\end{aligned}$$

and

$$V_n(y) = \varepsilon_n v_{\varepsilon_n}(y) + \frac{\delta\varepsilon_n}{2}|y|^2 - \frac{\langle B_2 Q_{2,N}(z_\alpha - y_\alpha), y\rangle}{\alpha} + \frac{|Q_{2,N}(y-y_\alpha)|^2_{2,-1}}{\alpha}$$

we see that

$$U_n(t,x,z) - V_n(y) - \frac{|P_{2,N}(z-y)|^2_{2,-1}}{2\alpha}$$

has a strict global maximum at $(t_\alpha, x_\alpha, z_\alpha, y_\alpha)$.

It now follows from the proof of Theorem 3.27 on [14] (see also Cor. 3.28 there) that there exist functions $\varphi_k, \psi_k \in C^2(H)$ such that $\varphi_k(z) = \varphi_k(P_{2,N}z), \psi_k(y) = \psi_k(P_{2,N}y)$, $\varphi_k, D\varphi_k, D^2\varphi_k, \psi_k$, $D\psi_k, D^2\psi_k$ are bounded and uniformly continuous, and such that

$$U_n(t,x,z)-\varphi_k(z) \quad \text{has a global maximum at some point } (\hat t_k,\hat x_k,\hat z_k),$$

$$V_n(y)-\psi_k(y) \quad \text{has a global minimum at some point } (\hat y_k),$$

and, as $k\to\infty$,

$$\begin{aligned}&(\hat t_k,\hat x_k,\hat z_k,U_n(\hat t_k,\hat x_k,\hat z_k),D\varphi_k(\hat z_k),D^2\varphi_k(\hat z_k))\\ &\to\left(t_\alpha,x_\alpha,z_\alpha,U_n(t_\alpha,x_\alpha,z_\alpha),\frac{B_2P_{2,N}(z_\alpha-y_\alpha)}{\alpha},X_N\right)\quad \text{in } \mathbb{R}\times H\times H\times\mathbb{R}\times H_{2,2}\times\mathcal{L}(H)\end{aligned}\tag{7.26}$$

$$\begin{aligned}&(\hat y_k,V_n(\hat y_k),D\psi_k(\hat y_k),D^2\psi_k(\hat y_k))\\ &\to\left(y_\alpha,V_n(y_\alpha),\frac{B_2P_{2,N}(z_\alpha-y_\alpha)}{\alpha},Y_N\right)\quad \text{in } H\times\mathbb{R}\times H_{2,2}\times\mathcal{L}(H),\end{aligned}\tag{7.27}$$

where $X_N=P_{2,N}X_NP_{2,N}, Y_N=P_{2,N}Y_NP_{2,N}$,

$$\begin{pmatrix} X & 0\\ 0 & -Y_N\end{pmatrix}\le\frac{3}{\alpha}\begin{pmatrix} B_2P_{2,N} & -B_2P_{2,N}\\ -B_2P_{2,N} & B_2P_{2,N}\end{pmatrix}.\tag{7.28}$$

(To be in the setup of Theorem 3.27 on [14] we can extend the number of variables of V_n by setting, say, $V_n(s,w,y)=V_n(y)-s^2-|w|^2$ so that we now have a strict maximum at $(t_\alpha,x_\alpha,z_\alpha,0,0,y_\alpha)$, carry out the proof and then restrict the conclusions to the statements about the functions restricted to the original variables.)

Using the definition of viscosity subsolution we now have $x_k\in H_{1,1}$ and

$$\begin{aligned}&\partial_t\psi(\hat t_k,\hat x_k)-\lambda\eta(\hat t_k)\frac{\theta'(|\hat x_k|)}{|\hat x_k|}|\hat x_k|_{1,1}^2-\delta\mu|\hat z_k|^2-\langle\hat x_k,A_1^*D\varphi(\hat t_k,\hat x_k)\rangle\\ &-\frac{1}{\varepsilon_n}\left\langle\hat z_k,A_2^*\left(D\varphi_k(\hat z_k)+\frac{B_2Q_{2,N}(z_\alpha-y_\alpha)}{\alpha}+\frac{2B_2Q_{2,N}(\hat z_k-z_\alpha)}{\alpha}\right)\right\rangle\\ &-F\left(\hat x_k,\hat z_k,D\psi(\hat t_k,\hat x_k),\delta\hat z_k+\frac{1}{\varepsilon_n}D\varphi_k(\hat z_k)+\frac{B_2Q_{2,N}(z_\alpha-y_\alpha)}{\varepsilon_n\alpha}+\frac{2B_2Q_{2,N}(\hat z_k-z_\alpha)}{\varepsilon_n\alpha},\right.\\ &\left.D^2\psi(\hat t_k,\hat x_k),\frac{1}{\varepsilon_n}D^2\varphi_k(\hat z_k)+\delta I+\frac{2B_2Q_{2,N}}{\varepsilon_n\alpha},0\right)\ge 0.\end{aligned}\tag{7.29}$$

Since (7.29) implies that, up to a subsequence, $\hat x_k\rightharpoonup x_\alpha$ in $H_{1,1}$, passing to the limit as $k\to\infty$ in (7.29) and using (3.12) and (7.26), we obtain

$$\begin{aligned}&\partial_t\psi(t_\alpha,x_\alpha)-\lambda\eta(t_\alpha)\frac{\theta'(|x_\alpha|)}{|x_\alpha|}|x_\alpha|_{1,1}^2-\delta\mu|z_\alpha|^2\\ &-\langle x_\alpha,A_1^*D\varphi(t_\alpha,x_\alpha)\rangle-\frac{1}{\varepsilon_n\alpha}\langle z_\alpha,A_2^*B_2(z_\alpha-y_\alpha)\rangle\\ &-F\left(x_\alpha,z_\alpha,D\psi(t_\alpha,x_\alpha),\delta z_\alpha+\frac{1}{\varepsilon_n\alpha}B_2(z_\alpha-y_\alpha),D^2\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n}X_N+\delta I,0\right)\ge\rho_1(N),\end{aligned}$$

where $\lim_{N\to\infty}\rho_1(N)=0$. It now follows

$$\begin{aligned}&\partial_t\psi(t_\alpha,x_\alpha)-\lambda\eta(t_\alpha)\frac{\theta'(|x_\alpha|)}{|x_\alpha|}|x_\alpha|_{1,1}^2-\langle x_\alpha,A_1^*D\varphi(t_\alpha,x_\alpha)\rangle-\frac{1}{\varepsilon_n\alpha}\langle z_\alpha,A_2^*B_2(z_\alpha-y_\alpha)\rangle\\&-F\left(x_\alpha,z_\alpha,D\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n\alpha}B_2(z_\alpha-y_\alpha),D^2\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n}X_N,0\right)\\&\geq\rho_1(N)-\delta(C(1+|z_\alpha|)-\mu|z_\alpha|^2)\end{aligned}\tag{7.30}$$

for some absolute constant $C>0$.

Using the definition of viscosity supersolution for the function v_{ε_n}, we obtain

$$\begin{aligned}&\varepsilon_nv_{\varepsilon_n}(\hat y_k)+\delta\mu|\hat y_k|^2+\frac{1}{\varepsilon_n}\left\langle\hat y_k,A_2^*\left(D\psi_k(\hat y_k)+\frac{1}{\alpha}B_2Q_{2,N}(\hat z_k-\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}(\hat y_k-y_\alpha)\right)\right\rangle\\&+F\left(\bar x,\hat y_k,D\psi(\bar t,\bar x),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_n\hat y_k+D\psi_k(\hat y_k)+\frac{1}{\alpha}B_2Q_{2,N}(\hat z_k-\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}(\hat y_k-y_\alpha)\right),\right.\\&\left.D^2\psi(\bar t,\bar x),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_nI+D^2\psi_k(\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}\right),0\right)-\overline{F}\geq0.\end{aligned}\tag{7.31}$$

Since the function v_{ε_n} is Lipschitz continuous on H we have

$$\left|-\delta\varepsilon_n\hat y_k+D\psi_k(\hat y_k)+\frac{1}{\alpha}B_2Q_{2,N}(\hat z_k-\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}(\hat y_k-y_\alpha)\right|\leq C\varepsilon_n$$

for some C independent of k,α and N. Therefore, by Assumption 3.1 (this is the only place that requires h to be independent of x),

$$\begin{aligned}&\left|F\left(\bar x,\hat y_k,D\psi(\bar t,\bar x),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_n\hat y_k+D\psi_k(\hat y_k)+\frac{1}{\alpha}B_2Q_{2,N}(\hat z_k-\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}(\hat y_k-y_\alpha)\right),\right.\right.\\&\left.D^2\psi(\bar t,\bar x),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_nI+D^2\psi_k(\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}\right),0\right)\\&-F\left(x_\alpha,\hat y_k,D\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_n\hat y_k+D\psi_k(\hat y_k)+\frac{1}{\alpha}B_2Q_{2,N}(\hat z_k-\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}(\hat y_k-y_\alpha)\right),\right.\\&\left.\left.D^2\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n}\left(-\delta\varepsilon_nI+D^2\psi_k(\hat y_k)-\frac{2}{\alpha}B_2Q_{2,N}\right),0\right)\right|\leq\rho_2(|\bar t-t_\alpha|+|\bar x-x_\alpha|)\end{aligned}$$

for some modulus ρ_2. Passing to the limit as $k\to\infty$ in (7.31) and using (3.12) and (7.27), we obtain

$$\begin{aligned}\varepsilon_nv_{\varepsilon_n}(y_\alpha)+\delta\mu|y_\alpha|^2+\frac{1}{\varepsilon_n\alpha}\langle y_\alpha,A_2^*B_2(z_\alpha-y_\alpha)\rangle+F\Bigg(x_\alpha,y_\alpha,D\psi(t_\alpha,x_\alpha),-\delta y_\alpha+\frac{1}{\varepsilon_n\alpha}B_2(z_\alpha-y_\alpha),\\D^2\psi(t_\alpha,x_\alpha),-\delta I+\frac{1}{\varepsilon_n}Y_N,0\Bigg)-\overline{F}\geq\rho_1(N)-\rho_2(|\bar t-t_\alpha|+|\bar x-x_\alpha|)\end{aligned}$$

and thus

$$\begin{aligned}&\varepsilon_n v_{\varepsilon_n}(y_\alpha)+\frac{1}{\varepsilon_n\alpha}\langle y_\alpha, A_2^*B_2(z_\alpha-y_\alpha)\rangle+F\bigg(x_\alpha,y_\alpha,D\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n\alpha}B_2(z_\alpha-y_\alpha),\\ &D^2\psi(t_\alpha,x_\alpha),\frac{1}{\varepsilon_n}Y_N,0\bigg)-\overline{F}\geq\rho_1(N)-\rho_2(|\bar t-t_\alpha|+|\bar x-x_\alpha|)-\delta(C(1+|y_\alpha|)-\mu|y_\alpha|^2).\end{aligned}\tag{7.32}$$

We now add (7.30) and (7.32), and use (3.3), (3.5), (3.6), (3.7), (3.8) (see [14], pages 234–236 for such arguments) to get

$$\begin{aligned}&\partial_t\psi(t_\alpha,x_\alpha)-\lambda\eta(t_\alpha)\frac{\theta'(|x_\alpha|)}{|x_\alpha|}|x_\alpha|^2_{1,1}-\langle x_\alpha,A_1^*D\varphi(t_\alpha,x_\alpha)\rangle\\ &+\varepsilon_n v_{\varepsilon_n}(y_\alpha)-\frac{1}{2\varepsilon_n\alpha}|z_\alpha-y_\alpha|^2+\frac{C}{\varepsilon_n\alpha}|z_\alpha-y_\alpha|^2_{2,-1}+\tilde\omega(|z_\alpha-y_\alpha|)-\overline{F}\\ &\geq\rho_1(N)-\rho_2(|\bar t-t_\alpha|+|\bar x-x_\alpha|)-\delta(C(2+|y_\alpha|+|z_\alpha|)-\mu(|y_\alpha|^2+|z_\alpha|^2))\end{aligned}\tag{7.33}$$

for some modulus $\tilde\omega$. By (7.25), the above implies that there exists a constant C_1, independent of n, such that for large α, $|y_\alpha|\leq C_1, |y_\alpha|\leq C_1, |x_\alpha|^2_{1,1}\leq C_1$, and hence, since $(t_\alpha,x_\alpha,z_\alpha,y_\alpha)\to(t_n,x_n,y_n,y_n)$, we must have $|y_n|\leq C_1$, $x_n\in H_{1,1}$ and $x_\alpha\rightharpoonup x_n$ in $H_{1,1}$, as $\alpha\to 0$. It thus follows by letting $N\to\infty$ and then $\alpha\to 0$ in the above inequality and using

$$\lim_{\alpha\to 0}\sup_{r>0}\left(\tilde\omega(r)-\frac{1}{2\varepsilon_n\alpha}r^2\right)=0$$

that

$$\begin{aligned}&\partial_t\psi(t_n,x_n)-\lambda\eta(t_n)\frac{\theta'(|x_n)}{|x_n|}|x_n|^2_{1,1}-\langle x_n,A_1^*D\varphi(t_n,x_n)\rangle\\ &+\varepsilon_n v_{\varepsilon_n}(y_n)-\overline{F}\geq-\rho_2(|\bar t-t_n|+|\bar x-x_n|)-\delta C_2\end{aligned}\tag{7.34}$$

for some constant C_2 independent of n and δ. Inequality (7.34) implies that $|x_n|_{1,1}\leq C$ for some constant C and hence, up to a subsequence, $t_n\to\tilde t, x_n\rightharpoonup\tilde x$ in $H_{1,1}$ for some $\tilde t$ and $\tilde x\in H_{1,1}$. Since B_1 is compact, this implies $x_n\to\tilde x$. Therefore

$$\overline{u}(\tilde t,\tilde x)\geq\limsup_{n\to\infty}u_{\varepsilon_n}(t_n,z_n,y_n)$$

and so

$$\begin{aligned}0&\geq(\overline{u}-\eta(\cdot)\theta(|\cdot|))^*(\tilde t,\tilde x)-\varphi(\tilde t,\tilde x)\geq\overline{u}(\tilde t,\tilde x)-\eta(\tilde t)\theta(|\tilde x|)-\varphi(\tilde t,\tilde x)\\ &\geq\limsup_{n\to\infty}(u_{\varepsilon_n}(t_n,z_n,y_n)-\varepsilon_n v_n(y_n)-\delta\varepsilon_n|y_n|^2-\eta(t_n)\theta(|x_n|)-\varphi(t_n,x_n))\geq 0.\end{aligned}$$

Thus $(\tilde t,\tilde x)=(\bar t,\bar x)$ and letting $n\to\infty$ and then $\delta\to 0$ in (7.34) gives

$$\partial_t\psi(\bar t,\bar x)-\lambda\eta(\bar t)\frac{\theta'(|\bar x)}{|\bar x|}|\bar x|^2_{1,1}-\langle\bar x,A_1^*D\varphi(\bar t,\bar x)\rangle-\overline{F}\geq 0,$$

which concludes the proof that $\overline{u}$ is a viscosity subsolution of (7.1).

Therefore, by comparison principle for (7.1), we now have $\overline{u} \le u \le \underline{u}$ on $(0,T] \times H$ and hence $\overline{u} = \underline{u} = u$. Let now $0 < T_1 < T, R > 0$ and $K \subset H$ be compact. If u_ε does not converge uniformly on $[T_1,T] \times K \times B(0,R)$, as $\varepsilon \to 0$ to u then there are sequences ε_n and $(t_n,x_n,y_n) \in [T_1,T] \times K \times B(0,R)$ such that $t_n \to t, x_n \to x$ for some t,x such that, say,

$$u_{\varepsilon_n}(t_n,x_n,y_n) \ge u(t_n,x_n) + \nu$$

for some $\nu > 0$. But then

$$\overline{u}(t,x) \ge \limsup_{n\to\infty} u_{\varepsilon_n}(t_n,x_n,y_n) \ge u(t,x) + \nu$$

which is impossible and thus the proof is complete. □

Remark 7.3. We point out that if the diffusion coefficient functions σ and h are both independent of x and y, assumption (3.6) is void and assumption (3.7) can be replaced by

$$|g(x_1,y_1,a) - g(x_2,y_2,a)| + |l(x_1,y_1,a) - l(x_1,y_2,a)| \le L(|x_1 - x_2| + |y_1 - y_2|) \tag{7.35}$$

for all $x_1,x_2,y_1,y_2 \in H, a \in \Lambda$. Then Theorems 4.8 and 5.1 hold without any changes, Lemma 6.1 is true if $|x_1 - x_2|_{1,-1}^2$ in (6.1) is replaced by $|x_1 - x_2|^2$, Lemma 6.2 and consequently Theorem 7.1 are true without any changes, and finally Theorem 7.2 holds. Thus, as far as the continuity of the coefficients, we cover the case considered by [16], however we have different assumptions on the operators A_1, A_2 and our diffusion coefficients must be Hilbert-Schmidt while [16] dealt with cylindrical Wiener processes and bounded constant diffusion coefficient operators.

Remark 7.4. It is easy to see by simple modifications of the proofs that all results of this paper except Theorem 7.2 are still true if the diffusion coefficient function h also depends on x and satisfies

$$\|h(x_1,y_1,a) - h(x_2,y_2,a)\|_{\mathcal{L}_2(K_0,H)} \le L(|x_1 - x_2|_{1,-1} + |y_1 - y_2|_{2,-1}) \quad \forall x_1,x_2,y_1,y_2 \in H, a \in \Lambda.$$

We finish the paper with some examples of problems where our theorems can be applied.

Example 7.5. We consider the singular perturbation and control problem for a system of stochastic parabolic partial differential equations. Let $\mathcal{O}$ be a bounded, smooth domain in $\mathbb{R}^n$ and $H = \Lambda := L^2(\mathcal{O})$. Let A_1, A_2 be the operators defined in (3.13) which satisfy (3.14) and the regularity conditions on the coefficients introduced there. As it was discussed after Remark 3.2, if the zero order coefficients e of both operators are positive and large enough, then both A_1, A_2 are maximal monotone and, if $B_1 = B_2 = \lambda_2(-\Delta)^{-1}$ for sufficiently large λ_2, they satisfy (3.2) and (3.3) for some $\lambda, c > 0$. Thus the norms $\|\cdot\|_{i,1}, \|\cdot\|_{i,-1}, i = 1,2$ are just the standard Sobolev and negative Sobolev norms. Obviously B_1, B_1 are compact. We also recall that, by Poincaré inequality, inequality (3.2) for A_2 implies (3.10) for some $\mu = \mu_1$. We now consider two cases.
Case 1. (Additive noise) Let $b_1, g_1, l_1 : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $f_1 : \mathbb{R} \to \mathbb{R}$ be bounded and continuous. We define b, g, l, f to be the Nemytskii operators. For $x, y, a \in L^2(\mathcal{O})$, we set

$$b(x,y,a)(\xi) = b_1(x(\xi),y(\xi),a(\xi)), \quad g(x,y,a)(\xi) = g_1(x(\xi),y(\xi),a(\xi))$$

$$l(x,y,a)(\xi) = l_1(x(\xi),y(\xi),a(\xi)), \quad f(x)(\xi) = f_1(x(\xi)).$$

It is easy to see that if b_1, g_1 are Lipschitz continuous in the first two variables, l_1 is Lipschitz continuous in the first variable and uniformly continuous in the second variable, and f_1 is uniformly continuous, then assumptions (3.5), (3.8) hold and (3.7) is replaced by (7.35). Suppose that $\sigma, h : \Lambda \to \mathcal{L}_2(K_0,H)$ are bounded and continuous.

Then assumption (3.6) is trivially satisfied. Moreover, if $\sigma : \Lambda \to \mathcal{L}(K, H)$ is bounded and Q is trace class in K, then assumption (3.9) holds. For instance we could take $K = H$ and $\sigma, h : \Lambda \to \mathcal{L}(H)$ be the multiplication operators defined by

$$[\sigma(a)z](\xi) = \sigma_1(a(\xi))z(\xi), \quad [h(a)z](\xi) = h_1(a(\xi))z(\xi)$$

for some continuous and bounded functions $\sigma_1, h_1 : \mathbb{R} \to \mathbb{R}$. Regarding assumption (3.4), it will be satisfied with $\mu = \mu_1 - \mu_2$ if there exists $\mu_2 < \mu_1$ such that for every $s, t_1, t_2, r \in \mathbb{R}$

$$(g_1(s, t_1, r) - g_1(s, t_2, r))(t_1 - t_2) \le \mu_2(t_1 - t_2)^2.$$

Therefore, in light of Remark 7.3, Theorem 7.2 holds in this case. We point out that in particular we can take $\sigma, h = 0$ in which case Theorem 7.2 holds for a deterministic singular perturbation and control problem.

Case 2. (Multiplicative finite dimensional noise) In the multiplicative noise case we face two issues. Nemytskii operators in general do not preserve Lipschitz continuity with respect to the operator and Hilbert-Schmidt norms and we have to deal with continuity with respect to weaker norms. To take care of the first issue we will assume that $K = K_0 = \mathbb{R}^m, Q = I \in S(\mathbb{R}^m)$ so W_Q is a standard Wiener process in $\mathbb{R}^m$. Regarding the second issue we recall that a bounded operator $P \in L(H)$ satisfies $|Px| \le C_P|B^{\frac{1}{2}}x|$ for some C_P and every $x \in H$, if and only if $P^*(H) \subset B^{\frac{1}{2}}(H)$, see *e.g.* [12], page 429, Proposition B.1.

Let P_1, P_2 be such operators. For instance they can be orthogonal projections onto the spaces spanned by finite numbers of eigenvectors of B. Let b_1, g_1, l_1, f_1 be as in Case 1 and let $\sigma_1, h_1 : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^m$ be bounded and continuous. We take b, f, l as in Case 1 and define versions of Nemytskii operators σ, h, l by setting, for $x, y, a \in L^2(\mathcal{O}), z \in \mathbb{R}^m$,

$$[\sigma(x, y, a)z](\xi) = \sigma_1([P_1x](\xi), [P_2y](\xi), a(\xi)) \cdot z, \quad [h(y, a)](\xi) = h_1([P_2y](\xi), a(\xi)) \cdot z,$$

$$g(x, y, a)(\xi) = g_1([P_1x](\xi), y(\xi), a(\xi)),$$

where $\cdot$ above is the dot product in $\mathbb{R}^m$. If b_1, g_1, l_1, f_1 satisfy the same assumptions as those in Case 1, then assumptions (3.5), (3.7), (3.8) and (3.9) (see Rem. 3.2) are satisfied. Regarding (3.6), in this case, by the definition of the Hilbert-Schmidt norm, we have

$$\begin{aligned}\|h(y_1, a) - h(y_2, a)\|^2_{\mathcal{L}_2(\mathbb{R}^m, H)} &= \int_{\mathcal{O}} |h_1([P_2y_1](\xi), a(\xi)) - h_1([P_2y_2](\xi), a(\xi))|^2 \mathrm{d}\xi \\ &\le L_{h_1}|P_2(y_1 - y_2)|^2 \le L_{h_1}C_{P_2}|y_1 - y_2|^2_{2,-1},\end{aligned}$$

where L_{h_1} is the Lipschitz constant of h_1 with respect to the first variable. Similar computation is also done for σ and thus (3.6) is satisfied.

Regarding assumption (3.4), we recall that if g_1 is as in Case 1, then we have for $x, a \in H, y_1, y_2 \in D(A_2^*)$,

$$-\langle A_2^*(y_1 - y_2), y_1 - y_2\rangle + \langle g(x, y_1, a) - g(x, y_2, a), y_1 - y_2\rangle \le -(\mu_1 - \mu_2)|y_1 - y_2|^2.$$

Thus, since by the properties of trace and the definition of the Hilbert-Schmidt norm,

$$\mathrm{Tr}\left((h(y_1, a) - h(y_2, a))Q^{\frac{1}{2}})((h(y_1, a) - h(y_2, a))Q^{\frac{1}{2}})^*\right) = \|h(y_1, a) - h(y_2, a)\|^2_{\mathcal{L}_2(K_0, H)},$$

assumption (3.4) will be satisfied with $\mu = \mu_1 - \mu_2 - \frac{1}{2} L_{h_1} \|P_2\|$ if $\mu_1 - \mu_2 > \frac{1}{2} L_{h_1} \|P_2\|$. Thus the combined linear and the nonlinear drift terms must be strongly monotone enough to compensate for the diffusion part, however this is required only if h depends on y.

Finally we mention that we could also take σ, h, g to be cylindrical functions. More precisely, if $e_1, ..., e_k$ are some elements of $D(B^{\frac{1}{2}})$, we take $\sigma_1 : \mathbb{R}^{2k+1} \to \mathbb{R}^m$, $h_1 : \mathbb{R}^{k+1} \to \mathbb{R}^m$, $g_1 : \mathbb{R}^{k+2} \to \mathbb{R}$ and define

$$\sigma(x, y, a)(\xi) = \sigma_1(\langle x, e_1\rangle, ..., \langle x, e_k\rangle, \langle y, e_1\rangle, ..., \langle y, e_k\rangle, a(\xi)),$$

$$h(y, a)(\xi) = h_1(\langle y, e_1\rangle, ..., \langle y, e_k\rangle, a(\xi)), \quad g(x, y, a)(\xi) = g_1(\langle x, e_1\rangle, ..., \langle x, e_k\rangle, y(\xi), a(\xi)).$$

Similar arguments to those used before show that if σ_1, h_1, g_1 are bounded, continuous and Lipschitz continuous in the variables involving x and y, then Assumption 3.1 is satisfied, provided $\mu_1 - \mu_2$ is large enough.

References

[1] O. Alvarez and M. Bardi, Viscosity solutions methods for singular perturbations in deterministic and stochastic control. *SIAM J. Control Optim.* **40** (2002) 1159–1188.

[2] O. Alvarez and M. Bardi, Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result. *Arch. Ration. Mech. Anal.* **170** (2003) 17–61.

[3] O. Alvarez and M. Bardi, Ergodicity, stabilization, and singular perturbations for Bellman-Isaacs equations. *Mem. Amer. Math. Soc.* **204** (2010) 1–88.

[4] M. Arisawa and P.-L. Lions, On ergodic stochastic control. *Commun. Partial Differ. Equ.* **23** (1998) 2187–2217.

[5] M. Bardi and A. Cesaroni, Optimal control with random parameters: a multiscale approach. *Eur. J. Control* **17** (2011) 30–45.

[6] M. Bardi, A. Cesaroni and L. Manca, Convergence by viscosity methods in multiscale financial models with stochastic volatility. *SIAM J. Financial Math.* **1** (2010) 230–265.

[7] A. Bensoussan, Perturbation methods in optimal control. Translated from the French by C. Tomson. *Wiley/Gauthier-Villars Series in Modern Applied Mathematics.* John Wiley & Sons, Ltd., Chichester; Gauthier-Villars, Montrouge (1988).

[8] V.S. Borkar and K. Suresh Kumar, Singular perturbations in risk-sensitive stochastic control. *SIAM J. Control Optim.* **48** (2010) 3675–3697.

[9] V.S. Borkar and V. Gaitsgory, Singular perturbations in ergodic control of diffusions. *SIAM J. Control Optim.* **46** (2007) 1562–1577.

[10] M.G. Crandall, H. Ishii and P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc.* **27** (1992) 1–67.

[11] M.G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms. *J. Funct. Anal.* **90** (1990) 237–283.

[12] G. Da Prato and J. Zabczyk, Stochastic equations in infinite dimensions. Second edition. *Encyclopedia of Mathematics and its Applications*, 152. Cambridge University Press, Cambridge (2014).

[13] T.D. Donchev and A.L. Dontchev, Singular perturbations in infinite-dimensional control systems. *SIAM J. Control Optim.* **42** (2003) 1795–1812.

[14] G. Fabbri, F. Gozzi and A. Święch, Stochastic optimal control in infinite dimension. Dynamic programming and HJB equations. With a contribution by Marco Fuhrman and Gianmario Tessitore. *Probability Theory and Stochastic Modelling,* 82. Springer, Cham (2017).

[15] D. Ghilli, Viscosity methods for large deviations estimates of multiscale stochastic processes. *ESAIM: COCV* **24** (2018) 605–637.

[16] G. Guatteri and G. Tessitore, Singular limit of BSDEs and optimal control of two scale stochastic systems in infinite dimensional spaces. To appear in: *Appl. Math. Optim.* (2019) https://doi.org/10.1007/s00245-019-09577-y.

[17] Y. Kabanov and S. Pergamenshchikov, Two-scale stochastic systems. Asymptotic analysis and control. Applications of Mathematics (New York), 49. *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin (2003).

[18] Y. Kabanov and W.J. Runggaldier, On control of two-scale stochastic systems with linear dynamics in the fast variables. *Math. Control Signals Syst.* **9** (1996) 107–122.

[19] D. Kelome and A. Święch, Perron's method and the method of relaxed limits for "unbounded" PDE in Hilbert spaces. *Studia Math.* **176** (2006) 249–277.

[20] H.J. Kushner, Weak convergence methods and singularly perturbed stochastic control and filtering problems. *Systems & Control: Foundations & Applications*, 3. Birkhäuser Boston, Inc., Boston, MA (1990).

[21] P.-L. Lions, Une inégalité pour les opérateurs elliptiques du second ordre. *Ann. Mat. Pura Appl.* **127** (1981) 1–11.

[22] P. Mannucci, C. Marchi and N. Tchou, Singular perturbations for a subelliptic operator. *ESAIM: COCV* **24** (2018) 1429–1451.

[23] D.S. Naidu, Singular perturbations and time scales in control theory and applications: an overview. Singularly perturbed dynamic systems in control technology. *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algor.* **9** (2002) 233–278.

[24] K. Suresh Kumar, Singular perturbations in stochastic ergodic control problems. *SIAM J. Control Optim.* **50** (2012) 3203–3223.
[25] J.T.F. Yang, *Singular perturbation of stochastic control and differential games*. Ph.D. thesis, The University of Sydney (2020).
[26] Y. Zhang, D.S. Naidu, C. Cai and Y. Zou, Singular perturbations and time scales in control theories and applications: an overview 2002–2012. *Int. J. Inf. Syst. Sci.* **9** (2014) 1–36.