



# Optimal control with random parameters: a multiscale approach

Martino Bardi, Annalisa Cesaroni

► **To cite this version:**

Martino Bardi, Annalisa Cesaroni. Optimal control with random parameters: a multiscale approach. European Journal of Control, Lavoisier, 2011, 17 (1), pp.30-45. <10.3166/ejc.17.30-45>. <hal-00664449>

**HAL Id: hal-00664449**

**<https://hal.inria.fr/hal-00664449>**

Submitted on 30 Jan 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimal control with random parameters: a multiscale approach\*

Martino Bardi<sup>†</sup> and Annalisa Cesaroni<sup>†</sup>

Revised version, June 2010

*Dedicated to Francis Clarke and Richard Vinter for their 60th birthday.*

## Abstract

We model the parameters of a control problem as an ergodic diffusion process evolving at a faster time scale than the state variables. We study the asymptotics as the speed of the parameters gets large. We prove the convergence of the value function to the solution of a limit Cauchy problem for a Hamilton-Jacobi equation whose Hamiltonian is a suitable average of the initial one. We give several examples where the effective Hamiltonian allows to define a limit control problem whose dynamics and payoff are linear or nonlinear averages of the initial data. This is therefore a constant-parameter approximation of the control problem with random entries. Our results hold if the fast random parameters are the only disturbances acting on the system, and then the limit system is deterministic, but also for dynamics affected by a white noise, and then the limit is a controlled diffusion.

**Keywords:** singular perturbations, viscosity solutions, deterministic control, stochastic control, asymptotic approximation, multiscale problems, well-posedness of control problems, sensitivity of control problems.

**AMS subject classification:** 35B25, 91B28, 93C70, 49L25.

## 1 Introduction

In all control problems the data (dynamical system, payoff functional,...) depend on several parameters that are often assumed constant, at least for short intervals of time, but may in fact change over time in a way that is usually unknown a priori. These parameters summarize the behaviour of all external un-modelled variables. A sequence of observations of these variables often looks like a sample of a stochastic process. One can take an average of them and use it as a constant parameter in the model. Alternatively, one can add some parameters to the state variables, assuming a dynamics consistent with the observed behavior. As an example,

---

\*Work partially supported by the Italian M.I.U.R. project "Viscosity, metric, and control theoretic methods for nonlinear partial differential equations".

<sup>†</sup>Dipartimento di Matematica P. e A., Università di Padova, via Trieste 63, 35121 Padova, Italy (bardi@math.unipd.it, acesar@math.unipd.it).

let us take a deterministic system in  $\mathbb{R}^n$  with state  $X_t$  and control  $u_t$  (with the notations of stochastic processes) and model the parameters  $Y_t \in \mathbb{R}^m$  as a given diffusion process:

$$\begin{cases} \dot{X}_t = f(X_t, Y_t, u_t), \\ dY_t = b(Y_t)dt + \sqrt{2\tau(Y_t)}dW_t, \end{cases} \quad (1)$$

where  $W_t$  is a Brownian motion. For the payoff functional, the most reasonable choice is taking the expectation with respect to the distribution of the process  $Y_t$ :

$$\mathbf{E} \left[ \int_t^T l(X_s, Y_s, u_s) ds + g(X_T, Y_T) \right]. \quad (2)$$

This model is more realistic than the one with constant parameters, but much harder to analyze, because the new system is a degenerate controlled diffusion and the increased dimension of the state space makes all computations more costly. The goal of this paper is to reconcile the two approaches by showing that, with a careful choice of the quantities to average, the constant-parameters model is a good approximation of the one with augmented state variables.

The main assumption we make is that the parameters  $Y_t$  are an ergodic process evolving on a faster time-scale than the "true" state variables  $X_t$  (see Section 2.4). This means that  $Y_t = \tilde{Y}_{t/\varepsilon}$  for a small  $\varepsilon > 0$  and the process  $\tilde{Y}_{\tilde{t}}$  has an invariant probability measure  $\mu$  such that

$$\lim_{T \rightarrow +\infty} \mathbf{E} \left[ \frac{1}{T} \int_0^T \phi(\tilde{Y}_{\tilde{t}}) d\tilde{t} \right] = \int_{\mathbb{R}^m} \phi(y) d\mu(y)$$

for all continuous  $\mu$ -integrable functions  $\phi$ , locally uniformly with respect to the initial position  $\tilde{Y}_0$ . Our result shows that in the limit as  $\varepsilon \rightarrow 0$  the problem of maximizing (2) for the  $n + m$  dimensional system (1) converges to a suitably  $\mu$ -averaged optimization problem for a deterministic  $n$ -dimensional system. Such *effective* problem is not always the same as for uncontrolled systems, because the limits  $\bar{f}, \bar{l}$  of the drift and running cost  $f, l$  can be different from the simple linear averages  $\int_{\mathbb{R}^m} f(x, y, u) d\mu(y)$ ,  $\int_{\mathbb{R}^m} l(x, y, u) d\mu(y)$ .

Before describing the result more precisely let us comment these assumptions. The ergodicity means that the process  $\tilde{Y}_\tau$  forgets its initial condition for large time and its distribution becomes stationary. The rescaled process  $Y_t$  satisfies a SDE of the form

$$dY_t = \frac{1}{\varepsilon} b(Y_t)dt + \sqrt{\frac{2}{\varepsilon}} \tau(Y_t) dW_t$$

and has the same properties on a finite time interval for small  $\varepsilon$ . Moreover its trajectories undergo rapid oscillations, therefore describing variables with a bursty behaviour. For these reasons the process  $Y_t$  was introduced to model some unknown parameters in financial mathematics since the 80s, with  $\varepsilon = 1$  first and then with  $\varepsilon$  small, see the books [26], [25] and the references therein. In that context the initial model for the state  $X_t$  is a diffusion process whose volatility is supposed to be a function of  $Y_t$ . The book by Fouque, Papanicolaou and Sircar [26] gives a nice survey of the empirical data supporting the stochastic volatility models, of the formal asymptotic expansion method for analyzing them, and of their applications to option pricing and hedging and to some optimization problems in financial markets. In these applications most authors choose for  $Y_t$  an Ornstein-Uhlenbeck process, that is also mean-reverting, Gaussian, and has an explicit formula for the density of the invariant measure  $\mu$ . See also [27, 28, 42, 7] for more recent developments and further references.

Another motivation for modeling the random parameters with a fast ergodic process is the following. Suppose  $\phi$  is a function of the parameters  $Y_t$  appearing in the model (e.g.,  $f, l, \dots$ ). A practitioner typically gets some historical values  $y_1, \dots, y_N$  of the parameters and then estimates  $\phi$  by the arithmetic mean of the observed data

$$\phi \approx \frac{1}{N} \sum_{i=1}^N \phi_i, \quad \phi_i := \phi(x, y_i, u).$$

Suppose the data  $y_i$  are samples of the process  $Y_t$  taken on a regular partition of the time interval  $[0, 1]$ , that is,  $y_i = Y_{i/N}$ . Then for large  $N$  and small  $\varepsilon$  the ergodicity of  $\tilde{Y}_\tau$  gives

$$\frac{1}{N} \sum_{i=1}^N \phi_i \approx \int_0^1 \phi(Y_t) dt = \varepsilon \int_0^{1/\varepsilon} \phi(\tilde{Y}_\tau) d\tau \approx \int_{\mathbb{R}^m} \phi(y) d\mu(y).$$

Once we have shown that the effective control problem obtained in the limit  $\varepsilon \rightarrow 0$  of the system (1) involves the average  $\int \phi d\mu$ , we can conclude that the arithmetic mean of the observed data is a good approximation of  $\phi(Y_t)$  in a constant-parameter model, provided there are many data and the parameters evolve fast enough.

We are left with the question: what are the right quantities to average? We give a simple answer: they are the terminal cost  $g$  and the Hamiltonian appearing in the Hamilton-Jacobi-Bellman equation, namely,

$$H(x, y, p) := \min_u \{-f(x, y, u) \cdot p - l(x, y, u)\}.$$

In fact, our main result states that the value function  $V^\varepsilon(t, x, y)$  of the maximization problem of the functional (2) for the system

$$\begin{cases} \dot{X}_s = f(X_s, Y_s, u_s), & X_t = x, \\ dY_s = \frac{1}{\varepsilon} b(Y_s) ds + \sqrt{\frac{2}{\varepsilon}} \tau(Y_s) dW_s, & Y_t = y, \end{cases} \quad (3)$$

converges locally uniformly to the viscosity solution  $V(t, x)$  of

$$-\frac{\partial V}{\partial t} + \int_{\mathbb{R}^m} H(x, y, D_x V) d\mu(y) = 0, \quad V(T, x) = \int_{\mathbb{R}^m} g(x, y) d\mu(y).$$

The *effective Hamiltonian*  $\bar{H}(x, p) := \int H(x, y, p) d\mu(y)$  is concave in the variables  $p$ , so it can be represented as a HJB Hamiltonian for suitable dynamics  $\bar{f}$  and running cost  $\bar{l}$ . These functions, together with the effective terminal cost  $\int g(x, y) d\mu(y)$ , define the effective (constant-parameters, deterministic) optimal control problem that approximates the stochastic model. However, there is not a general recipe for finding explicit formulas for  $\bar{f}$  and  $\bar{l}$ . In Section 4 we first give sufficient condition under which  $\bar{f}, \bar{l}$  are the linear averages of  $f, l$ , and then we show that different nonlinear averagings must be taken in a model from economics of Ramsey type (see [30], [17], [41]) and in one from advertising theory following Vidale and Wolfe (see [41], [33]).

The same effective Hamiltonian and limit control problem are also obtained if the equation for the state variables  $X_s$  in (3) is replaced by

$$dX_s = f(X_s, Y_s, u_s) ds + \varepsilon^\alpha dW_s$$

for some  $\alpha > 0$ . Therefore the effective control problem is the same if the dynamics is also affected by a small white noise, in addition to the fast oscillating stochastic parameters.

Our method is based on the HJB equation and uses the theory of viscosity solutions (see [6] and [20]). It allows us to treat a more general problem than (3), namely

$$\begin{cases} dX_t = f(X_t, Y_t, u_t)dt + \sqrt{2}\sigma^\varepsilon(X_t, Y_t, u_t)dW_t & X_{t_0} = x \in \mathbb{R}^n, \\ dY_t = \frac{1}{\varepsilon}b(X_t, Y_t)dt + \sqrt{\frac{2}{\varepsilon}}\tau(X_t, Y_t)dW_t & Y_{t_0} = y \in \mathbb{R}^m, \end{cases} \quad (4)$$

with  $\sigma^\varepsilon \rightarrow \sigma$  locally uniformly. Here the evolution of the  $X_s$  variables is disturbed by a white noise and the matrices  $\sigma^\varepsilon, \sigma$  satisfy only standard regularity and growth assumptions and can be degenerate, so that the previous cases are recovered by taking  $\sigma \equiv 0$  and either  $\sigma^\varepsilon \equiv 0$  or  $\sigma^\varepsilon = \varepsilon^\alpha \mathbf{I}$ . Moreover, the evolution of the  $Y_s$  variables is now no more decoupled, therefore describing the case when the parameters are influenced by the state  $X_s$ . This generality applies also to models where  $Y_s$  are not parameters but true state variables that evolve on a faster time scale than  $X_s$ , provided they are uncontrolled.

Our results fall within the domain of singular perturbations of diffusion processes and of control systems. There is a wide literature on this subject and its applications, see the recent survey papers [22, 38] and their large bibliographies. For results based on probabilistic methods we refer to the books [35, 34], the recent papers [39, 40, 13, 14], and the references therein. An approach based on the HJB equations started with [32, 10] and was developed within the theory of viscosity solutions by Alvarez and one of the authors in [1, 2, 3], see also [4, 5] for problems with an arbitrary number of scales. These methods originate in periodic homogenization theory [37, 24] and work nicely for fast variables restricted to a compact set, in particular the  $m$ -dimensional torus. The contribution of this paper is the treatment of unbounded fast variables  $Y_t$ , such as the classical Ornstein-Uhlenbeck process. In the companion paper [7] similar results were given for financial models with stochastic volatility, that have a special structure, null running cost, and dynamics of the fast variables independent of  $X_s$ . We use here some ergodic results of [7]. A convergence theorem in the unbounded setting was proved by Kushner [35] with probabilistic methods in the case of fast variables and controls appearing in a decoupled way, so that the effective system and cost are the linear averages of  $f$  and  $l$ . We get a variant of his theorem as a special case of ours in Section 4.1. Our results are also related to the theory of well-posedness of control problems [23] and to sensitivity analysis in optimization [18, 43, 12].

Finally let us mention that a different model of stochastic control problems with fast random oscillations was studied by Bensoussan and Blankenship [11]. It fits in the recent theory of homogenization of fully nonlinear elliptic PDEs in stationary ergodic media [16].

The paper is organized as follows. Section 2 presents the mathematical problem with the precise assumptions, the HJB equation, and the initial value problem satisfied by  $V^\varepsilon$ . In Section 3 we construct the effective Hamiltonian and terminal cost and prove the main result, Theorem 3.2, on the convergence of  $V^\varepsilon$  to the solution of the effective Cauchy problem. Section 4 is devoted to examples and applications. For the economic model of Section 4.2 we also discuss the convergence of the optimal feedback control for the problem with  $\varepsilon > 0$  to the one for the effective problem.

## 2 The two-scale stochastic control problem

### 2.1 The stochastic system

Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$  be a complete filtered probability space and let  $(W_t)_t$  be an  $\mathcal{F}_t$ -adapted standard  $r$ -dimensional Brownian motion. We consider the stochastic control system (4) where  $\varepsilon > 0$  and the coefficients satisfy the following standard conditions. We will use the symbols  $\mathbb{M}^{k,j}$  and  $\mathbb{S}^k$  to denote, respectively, the set of  $k \times j$  matrices and the set of  $k \times k$  symmetric matrices. For a given compact set  $U$ ,  $f : \mathbb{R}^n \times \mathbb{R}^m \times U \rightarrow \mathbb{R}^n$  and  $\sigma^\varepsilon : \mathbb{R}^n \times \mathbb{R}^m \times U \rightarrow \mathbb{M}^{n,r}$  are continuous functions, Lipschitz continuous in  $(x, y)$  uniformly w.r.t.  $u \in U$  and  $\varepsilon > 0$  and with linear growth in  $x$ , i.e.

$$\text{for some } C > 0 \quad |f(x, y, u)|, \|\sigma^\varepsilon(x, y, u)\| \leq C(1 + |x|) \quad \forall x, y, \forall \varepsilon > 0. \quad (5)$$

Moreover we assume that

$$\lim_{\varepsilon \rightarrow 0} \sigma^\varepsilon(x, y, u) = \sigma(x, y, u) \quad \text{locally uniformly,}$$

where  $\sigma : \mathbb{R}^n \times \mathbb{R}^m \times U \rightarrow \mathbb{M}^{n,r}$  satisfies the same conditions as  $\sigma^\varepsilon$ . The coefficients  $b : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  and  $\tau : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{M}^{m,r}$  are locally Lipschitz continuous functions with linear growth, i.e.

$$\text{for some } C > 0 \quad |b(x, y)|, \|\tau(x, y)\| \leq C(1 + |x| + |y|) \quad \text{for every } x, y. \quad (6)$$

Finally, the diffusion driving the fast variables  $Y_t$  is uniformly non degenerate, i.e.,

$$\exists \Lambda(y) > 0 \text{ such that } \xi \tau(x, y) \tau^T(x, y) \cdot \xi = |\xi \tau(x, y)|^2 \geq \Lambda(y) |\xi|^2 \quad \forall x, y, \xi. \quad (7)$$

We will not make any non-degeneracy assumption on the matrices  $\sigma^\varepsilon, \sigma$ , so the case  $\sigma \equiv 0$  is allowed.

### 2.2 The optimal control problem

We define the following pay off functional for a finite horizon optimal control problem associated to system (4)

$$J(t, x, y, u) := \mathbf{E} \left[ e^{\lambda(t-T)} g(X_T, Y_T) + \int_t^T l(X_s, Y_s, u_s) e^{\lambda(s-T)} ds \mid X_t = x, Y_t = y \right], \quad t \in [0, T],$$

where  $\mathbf{E}$  denotes the conditional expectation. The associated value function is

$$V^\varepsilon(t, x, y) := \sup_{u \in \mathcal{U}} J(t, x, y, u).$$

The discount factor is  $\lambda \geq 0$ . The utility function  $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  and the running cost  $l : \mathbb{R}^n \times \mathbb{R}^m \times U \rightarrow \mathbb{R}$  are continuous functions and satisfy

$$\exists K > 0 \text{ such that } \sup_{y \in \mathbb{R}^m} |g(x, y)|, \sup_{y \in \mathbb{R}^m, u \in U} |l(x, y, u)| \leq K(1 + |x|^2) \quad \forall x \in \mathbb{R}^n. \quad (8)$$

The set of admissible control functions  $\mathcal{U}$  is the standard one in stochastic control problems (see [25, Ch. IV, Definition 2]), i.e. it is the set of  $\mathcal{F}_t$ -progressively measurable processes taking values in  $U$ .

These conditions and those of the preceding Section will hold throughout the paper.

### 2.3 The HJB equation

It is well known that, under suitable growth conditions, the value function  $V^\varepsilon$  can be characterized as the unique continuous viscosity solution to an appropriate parabolic problem with terminal data.

The HJB equation associated via Dynamic Programming to the value function  $V^\varepsilon$  is

$$-V_t^\varepsilon + F^\varepsilon \left( x, y, V^\varepsilon, D_x V^\varepsilon, \frac{D_y V^\varepsilon}{\varepsilon}, D_{xx}^2 V^\varepsilon, \frac{D_{yy}^2 V^\varepsilon}{\varepsilon}, \frac{D_{xy}^2 V^\varepsilon}{\sqrt{\varepsilon}} \right) = 0 \quad \text{in } (0, T) \times \mathbb{R}^n \times \mathbb{R}^m, \quad (9)$$

complemented with the obvious terminal condition

$$V^\varepsilon(T, x, y) = g(x, y). \quad (10)$$

This is a fully nonlinear degenerate parabolic equation (strictly parabolic in the  $y$  variables by the assumption (7)).

The Hamiltonian  $F^\varepsilon : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{S}^n \times \mathbb{S}^m \times \mathbb{M}^{n,m} \rightarrow \mathbb{R}$  is defined as

$$F^\varepsilon(x, y, s, p, q, M, N, Z) := H^\varepsilon(x, y, p, M, Z) - \mathcal{L}(x, y, q, N) + \lambda s, \quad (11)$$

where

$$H^\varepsilon(x, y, p, M, Z) := \min_{u \in \mathcal{U}} \left\{ -\text{trace}(\sigma^\varepsilon(\sigma^\varepsilon)^T M) - f \cdot p - 2\text{trace}(\sigma^\varepsilon \tau^T Z^T) - l \right\} \quad (12)$$

with  $\sigma^\varepsilon$ ,  $f$  and  $l$  computed at  $(x, y, u)$ ,  $\tau = \tau(x, y)$ , and

$$\mathcal{L}(x, y, q, N) := b(x, y) \cdot q + \text{trace}(\tau(x, y)\tau^T(x, y)N). \quad (13)$$

We define also the Hamiltonian with  $\sigma^\varepsilon$  replaced by  $\sigma$  that will be useful in the following,

$$H(x, y, p, M, Z) := \min_{u \in \mathcal{U}} \left\{ -\text{trace}(\sigma\sigma^T M) - f \cdot p - 2\text{trace}(\sigma\tau^T Z^T) - l \right\}. \quad (14)$$

**Proposition 2.1.** *For any  $\varepsilon > 0$ , the function  $V^\varepsilon$  in Section 2.2 is the unique continuous viscosity solution to the Cauchy problem (9)-(10) with at most quadratic growth in  $x$  and  $y$ . Moreover there exist a constant  $K > 0$  independent of  $\varepsilon$  such that*

$$|V^\varepsilon(t, x, y)| \leq K(1 + |x|^2) \quad \text{for all } t \in [0, T], \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m. \quad (15)$$

*Proof.* This is a variant of a standard result (see [25] and the references therein) where we must take care of the unboundedness of the solution. So we just give a sketch of the proof.

Using definition of  $V^\varepsilon$  and assumption (8), it is easy to deduce that, for every  $\eta > 0$ , there exists  $u \in \mathcal{U}$  s.t.

$$|V^\varepsilon(t, x, y)| \leq \mathbf{E} \left[ K(1 + |X_T^u|^2) + K \int_t^T (1 + |X_s^u|^2) e^{-\lambda(T-s)} ds \middle| X_t = x, Y_t = y \right] + \eta.$$

Standard estimates on the second moment of the solution to (4) (see, for instance, [21, Lemma 3.1.] or [25, Appendix D]) and the boundedness of  $f$  and  $\sigma^\varepsilon$  with respect to  $y$  give that

$$\exists C > 0 \text{ s.t. } \sup_{0 \leq s \leq T} \mathbf{E}|X_s^u|^2 \leq (|x|^2 + CT)e^{CT}, \quad \forall u \in \mathcal{U}. \quad (16)$$

From this we get immediately (15) and this estimate in particular implies that the functions  $V^\varepsilon$  are locally equibounded.

We define the lower and upper semicontinuous envelope of  $V^\varepsilon$  as

$$V_*^\varepsilon(t, x, y) = \liminf_{(t', x', y') \rightarrow (t, x, y)} V^\varepsilon(t', x', y') \quad \text{and} \quad (V^\varepsilon)^*(t, x, y) = \limsup_{(t', x', y') \rightarrow (t, x, y)} V^\varepsilon(t', x', y').$$

By definition  $V_*^\varepsilon(t, x, y) \leq V^\varepsilon(t, x, y) \leq (V^\varepsilon)^*(t, x, y)$ . A standard argument in viscosity solution theory, based on the dynamic programming principle (see, e.g., [25, Ch. IV, Thm 7.1], [15, Thm 4.1]), gives that  $V_*^\varepsilon$  and  $(V^\varepsilon)^*$  are, respectively, a viscosity supersolution and a viscosity subsolution to (9), at every point  $(t, x, y) \in (0, T) \times \mathbb{R}^n \times \mathbb{R}^m$ .

Moreover  $\lim_{t \rightarrow T} V^\varepsilon(t, x, y) = g(x, y)$  locally uniformly in  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ . This result is well known and follows from (8), (16), and from the continuity in mean square of  $X_t, Y_t$ . For example, the same argument detailed in the proof of Proposition 3.1 in [7] can be repeated with minor changes.

Then the conclusion is obtained using a recent comparison result between sub and supersolutions to parabolic problems satisfying the quadratic growth condition

$$|V(t, x, y)| \leq K(1 + |x|^2 + |y|^2)$$

proved in Theorem 2.1 in [21]. Since  $V^\varepsilon$  satisfies (15) the result applies in our case, so  $(V^\varepsilon)^*(t, x, y) \leq V_*^\varepsilon(t, x, y)$ . Recalling the definition of semicontinuous envelopes, we get

$$(V^\varepsilon)^*(t, x, y) = V_*^\varepsilon(t, x, y) = V^\varepsilon(t, x, y) \quad \forall (t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m.$$

This implies that  $V^\varepsilon$  is a continuous viscosity solutions of (9), and, again by Theorem 2.1 in [21], that it is the unique solution with at most quadratic growth in  $x$  and  $y$ .  $\square$

## 2.4 Ergodicity of the fast variables

Consider the diffusion processes in  $\mathbb{R}^m$  obtained putting  $\varepsilon = 1$  in (4) and fixing  $x \in \mathbb{R}^n$

$$dY_t = b(x, Y_t)dt + \sqrt{2}\tau(x, Y_t)dW_t, \quad (17)$$

called the *fast subsystems*. To recall the dependance on the parameter  $x$ , we will denote the process in (17) as  $Y^x$ . Observe that its infinitesimal generator is  $\mathcal{L}_x w := \mathcal{L}(x, y, D_y w, D_{yy}^2 w)$ , with  $\mathcal{L}$  defined by (13).

Throughout the paper, we will assume the following condition:  
 $\forall x \in \mathbb{R}^n$  there exists  $w \in \mathcal{C}(\mathbb{R}^m)$ ,  $k > 0, R_0 > 0$  such that

$$-\mathcal{L}_x w \geq k \text{ for } |y| > R_0 \text{ in viscosity sense, and } w(y) \rightarrow +\infty \text{ as } |y| \rightarrow +\infty. \quad (18)$$

This condition is reminiscent of other similar conditions about ergodicity of diffusion processes in the whole space, see for example [29], [10], [36], [14], [13]. Lions and Musiela [36] also state that (18) is indeed equivalent to the ergodicity of the process in (17) and to the classical Lyapunov-type condition of Hasminskii [29].

**Remark 2.1.** Condition (18) can be interpreted as a weak Lyapunov condition for the process (17) relative to the set  $\{|y| \leq R_0\}$ . Indeed, a Lyapunov function for the system (17) relative to a compact invariant set  $K$  is a continuous, positive definite function  $L$  such that  $L(y) = 0$  if and only if  $y \in K$ , the sublevel sets  $\{y \mid L(y) \leq k\}$  are compact and  $-\mathcal{L}_x L(y) = l(y)$  in  $\mathbb{R}^m$ , where  $l$  is a continuous function with  $l = 0$  on  $K$  and  $l > 0$  outside. For more details see [29].



**Example 2.1.** Condition (18) is satisfied if, for every  $x \in \mathbb{R}^n$ ,

$$\limsup_{|y| \rightarrow +\infty} [b(x, y) \cdot y + \text{trace}(\tau(x, y)\tau^T(x, y))] < 0.$$

In this case it is immediate to check (18) by choosing  $w(y) = |y|^2$  and  $R_0$  sufficiently big. A typical model process which satisfies the previous condition is the Ornstein-Uhlenbeck process with equation

$$dY_t = (m(x) - Y_t)dt + \sqrt{2}\tau(x)dW_t.$$

Pardoux and Veretennikov [39, 40] assume  $\tau\tau^T$  bounded and  $\lim_{|y| \rightarrow +\infty} \sup_x b(x, y) \cdot y = -\infty$ , and call it recurrence condition.

In the following we give two results from [7] saying that, under (18) and the standing assumptions in Section 2.1, the process  $Y^x$  in (17) is ergodic, in the sense that it has a unique invariant measure, and a Liouville property holds. Moreover, we will discuss the regularity of the invariant measure w.r.t. the frozen variable  $x$ . The Liouville property replaces the standard strong maximum principle of the periodic case and is the key ingredient for extending some results of [3] to the non-periodic setting. The proof is in [7], Lemma 4.1.

**Lemma 2.1.** *For  $x \in \mathbb{R}^n$  fixed, consider the problem*

$$-\mathcal{L}_x V = -\mathcal{L}(x, y, DV(y), D^2V(y)) = 0 \quad y \in \mathbb{R}^m. \quad (19)$$

Then

- i) every bounded viscosity subsolution to (19) is constant;*
- ii) every bounded viscosity supersolution to (19) is constant.*

Next we state the existence and uniqueness of an invariant measure, see Proposition 4.2 in [7] (see also chapter IV in [29]).

**Proposition 2.2.** *Under the standing assumptions, for every  $x \in \mathbb{R}^n$ , there exists a unique invariant probability measure  $\mu_x$  on  $\mathbb{R}^m$  for the process  $Y^x$ .*

**Example 2.2.** For the multi-dimensional Ornstein-Uhlenbeck process

$$dY_t = (m(x) - Y_t)dt + \sqrt{2}\tau(x)dW_t.$$

the invariant measure has the explicit expression (see e.g. Section 4.1.2 in [19], see also Section 3.2.3 in [26])

$$d\mu_x(y) = \frac{1}{(2\pi)^{m/2} \det(\tau\tau^T(x))} \exp \left[ -\frac{1}{2}(\tau\tau^T(x))^{-1}(y - m(x)) \cdot (y - m(x)) \right] dy \quad \forall x.$$

Finally, the last result is about the Lipschitz regularity of  $\mu_x$  with respect to  $x$ . This result will be a key ingredient to prove locally Lipschitz regularity in  $x$  of the effective Hamiltonian (24). Obviously, when  $b, \tau$  do not depend on  $x$  this regularity property is trivially satisfied.

**Proposition 2.3.** *Besides the standing assumptions, assume that*

$$b \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^m) \quad \text{and} \quad \tau \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{M}^{m,r}) \quad (20)$$

*with all their derivatives bounded and Hölder continuous in  $y$  uniformly in  $x$ . Then the invariant measure  $\mu_x$  of the process  $Y^x$  has a density  $\varphi_x(y)$  and there exist  $k > 1$ ,  $C > 0$ , such that*

$$|\varphi_{x_1}(y) - \varphi_{x_2}(y)| \leq C \frac{1}{1 + |y|^k} |x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^n, \quad y \in \mathbb{R}^m.$$

*Proof.* For the proof we refer to Theorem 6 in [40], where such result is proved by PDE methods. The idea is to consider the adjoint operator to  $\mathcal{L}_x$  and to study the dependence of the solutions of the adjoint equation on the parameter  $x$ . The main tools are non trivial estimates on the fundamental solutions of nondegenerate second order parabolic PDEs.  $\square$

**Remark 2.2.** In [13] there is an analogous result (Prop. 5.2) about the Lipschitz regularity w.r.t.  $x$  of the effective system obtained by the average of a singularly perturbed stochastic system such as (4), under different assumptions and using mainly stochastic control methods. The authors replace regularity conditions such as (20) with appropriate estimates on the trajectories of the system, such as:

$$\exists C > 0 \text{ s.t. } \mathbf{E}|Y_t^{x_1} - Y_t^{x_2}| \leq C|x_1 - x_2| \quad \forall t \geq 0$$

where  $Y^{x_i}$  is the process in (17) with  $x = x_i$  and  $Y_0 = y$ . The authors state that this estimate can be obtained by an exponential stability condition on the process  $Y$  such as

$$\exists a, b > 0 \text{ s.t. } \mathbf{E}|(Y')_t^x - (Y'')_t^x| \leq ae^{-bt}|y' - y''|, \quad \forall t \geq 0, \quad \forall x,$$

where  $(Y')^x$  and  $(Y'')^x$  both satisfy (17) with initial condition respectively  $(Y')_0^x = y'$  and  $(Y'')_0^x = y''$ .

### 3 The convergence result

This section is devoted to the main result of the paper, namely, the convergence theorem for the singular perturbation problem. First of all, we will construct the effective Hamiltonian  $\overline{H}$  and the effective terminal data  $\overline{g}$ . Then we will prove in Theorem 3.2 that the value function  $V^\varepsilon(t, x, y)$ , solution to (9), converges locally uniformly, as  $\varepsilon \rightarrow 0$ , to a function  $V(t, x)$  which will be characterized as the unique solution of the Cauchy problem

$$\begin{cases} -V_t + \overline{H}(x, D_x V, D_{xx}^2 V) + \lambda V(x) = 0 & \text{in } (0, T) \times \mathbb{R}^n \\ V(T, x) = \overline{g}(x) & \text{in } \mathbb{R}^n. \end{cases} \quad (21)$$

#### 3.1 The effective Hamiltonian and initial data

Section 2.4 contains the main tools to define the candidate limit Cauchy problem of the singularly perturbed problem (9) as  $\varepsilon \rightarrow 0$ . We start showing the existence of an effective

Hamiltonian giving the limit PDE. In principle, for each  $(\bar{x}, \bar{p}, \bar{P})$  one expects the effective Hamiltonian  $\bar{H}(\bar{x}, \bar{p}, \bar{P})$  to be the unique constant  $c \in \mathbb{R}$  such that the *cell problem*

$$-\mathcal{L}(\bar{x}, y, D\chi, D^2\chi) + H(\bar{x}, y, \bar{p}, \bar{P}, 0) = c \quad \text{in } \mathbb{R}^m, \quad (22)$$

where  $H$  is defined in (14), has a viscosity solution  $\chi$ , called corrector (see [37], [24], [1]). Actually, for our approach, it is sufficient to consider, as in [2], a  $\delta$ -cell problem

$$\delta w_\delta - \mathcal{L}(\bar{x}, y, Dw_\delta, D^2w_\delta) + H(\bar{x}, y, \bar{p}, \bar{P}, 0) = 0 \quad \text{in } \mathbb{R}^m, \quad (23)$$

whose solution  $w_\delta$  is called *approximate corrector*. The next result states that  $\delta w_\delta$  converges to  $-\bar{H}$  and it is smooth.

**Theorem 3.1.** *Under the standing assumptions in Sections 2.1, 2.2 and 2.4, for any fixed  $(\bar{x}, \bar{p}, \bar{P})$  and  $\delta > 0$  there exists a solution  $w_\delta = w_{\delta; \bar{x}, \bar{p}, \bar{P}}(y)$  in  $\mathcal{C}^2(\mathbb{R}^m)$  of (23) such that*

$$-\lim_{\delta \rightarrow 0} \delta w_\delta = \bar{H}(\bar{x}, \bar{p}, \bar{P}) := \int_{\mathbb{R}^m} H(\bar{x}, y, \bar{p}, \bar{P}, 0) d\mu_{\bar{x}}(y) \quad \text{locally uniformly in } \mathbb{R}^m, \quad (24)$$

where  $\mu_{\bar{x}}$  is the invariant probability measure on  $\mathbb{R}^m$  for the process  $Y^{\bar{x}}$ .

The proof is given in [7], Theorem 4.3.

We define now the effective terminal value for the limit as  $\varepsilon \rightarrow 0$  of the singular perturbation problem (9). We fix  $\bar{x}$  and consider the following Cauchy initial problem:

$$\begin{cases} w_t - \mathcal{L}(\bar{x}, y, Dw, D^2w) = 0 & \text{in } (0, +\infty) \times \mathbb{R}^m \\ w(0, y) = g(\bar{x}, y), \end{cases} \quad (25)$$

where  $g$  satisfies assumption (8).

**Proposition 3.1.** *Under the standing assumptions in Sections 2.1, 2.2 and 2.4, for every  $\bar{x}$  there exists a unique bounded classical solution  $w(\cdot, \cdot; \bar{x})$  to (25) and*

$$\lim_{t \rightarrow +\infty} w(t, y; \bar{x}) = \int_{\mathbb{R}^m} g(\bar{x}, y) d\mu_{\bar{x}}(y) =: \bar{g}(\bar{x}) \quad \text{locally uniformly in } y. \quad (26)$$

Moreover  $\bar{g}$  is continuous and satisfy a quadratic growth condition as (8).

*Proof.* The proof can be found in [7] Proposition 4.4. The regularity of  $\bar{g}$  and the growth condition can be obtained, using the definition of  $\bar{g}$ , from condition (8) and the regularity of the invariant measure stated in Proposition 2.3.  $\square$

We say that the *Comparison Principle* holds for a Cauchy problem as (21) if given  $U \in USC([0, T] \times \mathbb{R}^n)$  and  $V \in LSC([0, T] \times \mathbb{R}^n)$ , respectively, a subsolution and a supersolution to the PDE and such that

$$\begin{aligned} U(T, x) &\leq V(T, x) && \text{for every } x, \\ \exists C > 0 \text{ s.t. } & |U(t, x)|, |V(t, x)| \leq C(1 + |x|^2) && \text{for every } t, x, \end{aligned}$$

then  $U(t, x) \leq V(t, x)$  in  $[0, T] \times \mathbb{R}^n$ .

**Proposition 3.2.** *Under the standing assumptions in Sections 2.1, 2.2 and 2.4, the Comparison Principle holds for the effective Cauchy problem (21) with  $\overline{H}$  and  $\overline{g}$  given by (24) and (26) if at least one of the following five sets of assumptions is verified:*

- (i)  $b, \tau$  are independent of  $x$ ;
- (ii)  $\sigma^\varepsilon \rightarrow \sigma \equiv 0$  and  $b, \tau$  satisfy the regularity assumptions (20);
- (iii) the fast subsystem is Ornstein-Uhlenbeck, i.e.,  $b(x, y) = m(x) - y$  and  $\tau(x, y) = \tau(x)$ , with  $m \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ ,  $\tau \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{M}^{m,r})$ , and  $\tau(x)\tau^T(x) \geq \tau_0 \mathbf{I}$  with  $\tau_0 > 0$ ;
- (iv)  $\sigma^\varepsilon \rightarrow \sigma(x, u)$  (independent of  $y$ ) and  $b, \tau$  satisfy the regularity assumptions (20);
- (v)  $\sigma^\varepsilon \rightarrow \sigma(x, y)$  (independent of  $u$ ), there exists a Lipschitz  $\overline{\sigma} : \mathbb{R}^n \rightarrow \mathbb{M}^{n,r}$  such that

$$\overline{\sigma}(x)\overline{\sigma}(x)^T = \int_{\mathbb{R}^m} \sigma(x, y)\sigma^T(x, y) d\mu_x(y),$$

$f$  and  $l$  are bounded,  $b, \tau$  satisfy the regularity assumptions (20), for some  $C > 0$   $g$  satisfies

$$|g(x_1, y) - g(x_2, y)| \leq C(1 + |x_1| + |x_2|)|x_1 - x_2| \quad \forall x_1, x_2, y,$$

and either  $g$  is independent of  $y$ , or it has at most linear growth in  $x$ :

$$|g(x, y)| \leq C(1 + |x|).$$

*Proof.* In the case (i) the effective Hamiltonian is the average with respect to a measure  $\mu(y)$  of a Hamiltonian satisfying the structural assumptions of [21] for the Comparison Principle, uniformly in  $y$ . It is immediate to check that the same proof as in [21, Thm 2.1] holds for the Cauchy problem with averaged Hamiltonian as well. Similarly, it is easy to check that the same arguments work also in case (iv). Finally, also in case (iii) the Comparison Principle can be obtained following the proof of [21, Thm 2.1]. Indeed the main point in this case is to check that for every  $R > 0$  there exists some  $C_R > 0$  such that, for every  $u \in U$ ,

$$\int_{\mathbb{R}^m} \left| \sigma(x_1, y, u)\sqrt{\varphi_{x_1}(y)} - \sigma(x_2, y, u)\sqrt{\varphi_{x_2}(y)} \right|^2 dy \leq C_R|x_1 - x_2|^2, \quad \forall x_1, x_2 \in B(0, R).$$

This is satisfied when, e.g.,

$$\int_{\mathbb{R}^m} \frac{|D_x \varphi_x(y)|^2}{\varphi_x(y)} dy \leq C_R, \quad \forall x \in B(0, R).$$

This is the case of the Ornstein-Uhlenbeck process under the assumptions in (iii), as it is possible to check with straightforward computations on the explicit formula given in Example 2.2.

Case (ii) is actually a particular case of (iv), but we state it separately because the effective Hamiltonian  $\overline{H}$  is of the first order and so the proof is much easier. Using the standing assumptions and Proposition 2.3 one checks that  $\overline{H}$  satisfies the following structural conditions: there exists  $C > 0$  such that

$$|\overline{H}(x, p_1) - \overline{H}(x, p_2)| \leq C(1 + |x|)|p_1 - p_2|, \quad \forall x, p_1, p_2 \in \mathbb{R}^n, \quad (27)$$

and for every  $R > 0$  there exists a continuity modulus  $\omega_R$  s.t.

$$|\overline{H}(x_1, p) - \overline{H}(x_2, p)| \leq \omega_R(|x_1 - x_2|(1 + |p|)), \quad \forall x_1, x_2 \in B(0, R), p \in \mathbb{R}^n. \quad (28)$$

Under these conditions, a classical Comparison Principle holds, see Theorem 2.5 in [31] or Theorem V.3.15 in [6] for the case of continuous sub and supersolutions, . The adaptation to the case of semicontinuous sub and supersolutions is straightforward, see [6, Exercise V.5.17].

In case (v), it is easy to check, using the standing assumptions and Proposition 2.3, that  $\overline{H}$  is the semilinear operator  $\overline{H}(x, p, P) = -\text{trace}(\overline{\sigma}\overline{\sigma}^t P) + K(x, p)$  and  $K$  satisfies the structural conditions (27), (28) with  $\omega_R \equiv \omega$  for all  $R > 0$ . Moreover, standard calculations show that the effective terminal data  $\overline{g}$  satisfies for some  $\overline{C}$

$$|\overline{g}(x_1) - \overline{g}(x_2)| \leq \overline{C}(1 + |x_1| + |x_2|)|x_1 - x_2| \quad \forall x_1, x_2.$$

Therefore the Comparison Principle follows from Theorem 2.1 in [9].  $\square$

### 3.2 The convergence theorem

We will assume in the following that the Comparison Principle holds for the effective Cauchy problem (21). We refer to Proposition 3.2 for a list of sufficient conditions ensuring its validity.

**Theorem 3.2.** *We assume, besides the standing assumptions in Sections 2.1, 2.2 and 2.4, that (21) satisfies the Comparison Principle, and that either  $g = g(x)$  is independent of  $y$  or that the coefficients  $b, \tau$  are independent of  $x$ . Then the solution  $V^\varepsilon$  to (9) converges uniformly on compact subsets of  $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  to the unique continuous viscosity solution to the limit problem (21) satisfying a quadratic growth condition in  $x$ , i. e.,*

$$\exists K > 0 \quad \text{s.t.} \quad |V(t, x)| \leq K(1 + |x|^2) \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n. \quad (29)$$

*In particular, if  $g$  is independent of  $y$ , then the convergence is uniform on compact subsets of  $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  and  $\overline{g} = g$ .*

*Proof.* The proof is divided in several steps.

#### Step 1 (Relaxed semilimits ).

Recall that by (15) the functions  $V^\varepsilon$  are locally equibounded in  $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m$ , uniformly in  $\varepsilon$ . We define the half-relaxed semilimits in  $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m$  (see [6, Ch V]):

$$\underline{V}(t, x, y) = \liminf_{\varepsilon \rightarrow 0, t' \rightarrow t, x' \rightarrow x, y' \rightarrow y} V^\varepsilon(t', x', y') \quad \overline{V}(t, x, y) = \limsup_{\varepsilon \rightarrow 0, t' \rightarrow t, x' \rightarrow x, y' \rightarrow y} V^\varepsilon(t', x', y')$$

for  $t < T$ ,  $x \in \mathbb{R}^n$  and  $y \in \mathbb{R}^m$ ,

$$\underline{V}(T, x, y) = \liminf_{t' \rightarrow T^-, x' \rightarrow x, y' \rightarrow y} \underline{V}(t', x', y') \quad \overline{V}(T, x, y) = \limsup_{t' \rightarrow T^-, x' \rightarrow x, y' \rightarrow y} \overline{V}(t', x', y').$$

It is immediate to get by definitions that the estimates (15) hold also for  $\overline{V}$  and  $\underline{V}$ . This means that

$$|\underline{V}(t, x, y)|, |\overline{V}(t, x, y)| \leq K(1 + |x|^2) \quad \text{for all } t \in [0, T], x \in \mathbb{R}^n, y \in \mathbb{R}^m. \quad (30)$$

**Step 2 ( $\bar{V}$ ,  $\underline{V}$  do not depend on  $y$ ).**

We check that  $\bar{V}(t, x, y), \underline{V}(t, x, y)$  do not depend on  $y$ , for every  $t \in [0, T]$  and  $x \in \mathbb{R}^n$ . We claim that  $\bar{V}(t, x, y)$  (resp.,  $\underline{V}(t, x, y)$ ) is, for every  $t \in (0, T)$  and  $x \in \mathbb{R}^n$ , a viscosity subsolution (resp., supersolution) to

$$-\mathcal{L}(x, y, D_y V, D_{yy}^2 V) = 0 \quad \text{in } \mathbb{R}^m \quad (31)$$

where  $\mathcal{L}$  is the differential operator defined in (13). If the claim is true, we can use Lemma 2.1, since  $\bar{V}, \underline{V}$  are bounded in  $y$  according to estimates (30), to conclude that the functions  $y \rightarrow \bar{V}(t, x, y), y \rightarrow \underline{V}(t, x, y)$  are constants for every  $(t, x) \in (0, T) \times \mathbb{R}^n$ . Finally, using the definition it is immediate to see that this implies that also  $\bar{V}(T, x, y)$  and  $\underline{V}(T, x, y)$  do not depend on  $y$ . We prove the claim only for  $\bar{V}$ , since the other case is completely analogous.

First of all we show that the function  $\bar{V}(t, x, y)$  is a viscosity subsolution to (31). To do this, we fix a point  $(\bar{t}, \bar{x}, \bar{y})$  and a smooth function  $\psi$  such that  $\bar{V} - \psi$  has a maximum at  $(\bar{t}, \bar{x}, \bar{y})$  in  $\bar{B} = \bar{B}((\bar{t}, \bar{x}, \bar{y}), r)$ , for some fixed  $r > 0$ . Using the definition of weak relaxed semilimits it is possible to prove (see [6, Lemma V.1.6]) that there exists  $\varepsilon_n \rightarrow 0$  and  $\bar{B} \ni (t_n, x_n, y_n) \rightarrow (\bar{t}, \bar{x}, \bar{y})$  maxima for  $V^{\varepsilon_n} - \psi$  in  $\bar{B}$  such that  $V^{\varepsilon_n}(t_n, x_n, y_n) \rightarrow \bar{V}(\bar{t}, \bar{x}, \bar{y})$ . Therefore, recalling that  $V^\varepsilon$  is a subsolution to (9), we get

$$-\psi_t + H\left(x_n, y_n, D_x \psi, D_{xx}^2 \psi, \frac{1}{\sqrt{\varepsilon_n}} D_{xy}^2 \psi\right) - \frac{1}{\varepsilon_n} \mathcal{L}(x_n, y_n, D_y \psi, D_{yy}^2 \psi) + \lambda V^{\varepsilon_n}(t_n, x_n, y_n) \leq 0,$$

where all the derivatives of  $\psi$  are computed resp. in  $(t_n, x_n, y_n)$ . This implies

$$-\mathcal{L}(x_n, y_n, D_y \psi, D_{yy}^2 \psi) \leq \varepsilon_n \left[ \psi_t - H\left(x_n, y_n, D_x \psi, D_{xx}^2 \psi, \frac{1}{\sqrt{\varepsilon_n}} D_{xy}^2 \psi\right) - \lambda V^{\varepsilon_n}(t_n, x_n, y_n) \right].$$

We observe that the part between brackets in the right-hand side of the previous is uniformly bounded with respect to  $n$  in  $\bar{B}$  and using the regularity properties of  $\psi$  and of the coefficients in the equation we get, as  $\varepsilon_n \rightarrow 0$ , the desired conclusion.

We show now that if  $\bar{V}(t, x, y)$  is a subsolution to (31), then for every fixed  $(\bar{t}, \bar{x})$  the function  $y \mapsto \bar{V}(\bar{t}, \bar{x}, y)$  is a subsolution to (31), which was our claim. To do this, we fix  $\bar{y}$  and a smooth function  $\phi$  such that  $\bar{V}(\bar{t}, \bar{x}, \cdot) - \phi$  has a strict local maximum at  $\bar{y}$  in  $B(\bar{y}, \delta)$  and such that  $\phi(y) \geq 1$  for all  $y \in B(\bar{y}, \delta)$ . We define, for  $\eta > 0$ ,  $\phi_\eta(t, x, y) = \phi(y) \left(1 + \frac{|x - \bar{x}|^2 + |t - \bar{t}|^2}{\eta}\right)$  and we consider  $(t_\eta, x_\eta, y_\eta)$  a maximum point of  $\bar{V} - \phi_\eta$  in  $B((\bar{t}, \bar{x}, \bar{y}), \delta)$ . Repeating the same argument as in [6, Lemma II.5.17], it is possible to prove, eventually passing to subsequences, that, as  $\eta \rightarrow 0$ ,  $(t_\eta, x_\eta, y_\eta) \rightarrow (\bar{t}, \bar{x}, \bar{y})$  and  $K_\eta := \left(1 + \frac{|x_\eta - \bar{x}|^2 + |t_\eta - \bar{t}|^2}{\eta}\right) \rightarrow K > 0$ . Moreover, using the fact that  $\bar{V}$  is a subsolution to (31), we get  $-\mathcal{L}(x_\eta, y_\eta, K_\eta D \phi(y_\eta), K_\eta D^2 \phi(y_\eta)) \leq 0$ , which gives, using the linearity of  $\mathcal{L}$  and passing to the limit as  $\eta \rightarrow 0$ , the desired conclusion  $-\mathcal{L}(\bar{x}, \bar{y}, D \phi(\bar{y}), D^2 \phi(\bar{y})) \leq 0$ .

**Step 3 ( $\bar{V}$  and  $\underline{V}$  are sub and supersolutions of the limit PDE).**

First we claim that  $\bar{V}$  and  $\underline{V}$  are sub and supersolution to the PDE in (21) in  $(0, T) \times \mathbb{R}^n$ . We prove the claim only for  $\bar{V}$  since the other case is completely analogous. The proof adapts the perturbed test function method introduced in [24] for the periodic setting. We fix  $(\bar{t}, \bar{x}) \in ((0, T) \times \mathbb{R}^n)$  and we show that  $\bar{V}$  is a viscosity subsolution at  $(\bar{t}, \bar{x})$  of the limit problem. This means that if  $\psi$  is a smooth function such that  $\psi(\bar{t}, \bar{x}) = \bar{V}(\bar{t}, \bar{x})$  and  $\bar{V} - \psi$  has a maximum at  $(\bar{t}, \bar{x})$  then

$$-\psi_t(\bar{t}, \bar{x}) + \bar{H}(\bar{x}, D_x \psi(\bar{t}, \bar{x}), D_{xx}^2 \psi(\bar{t}, \bar{x})) + \lambda \bar{V}(\bar{t}, \bar{x}) \leq 0. \quad (32)$$

Without loss of generality we assume that the maximum is strict in  $B((\bar{t}, \bar{x}), r)$  and that  $0 < \bar{t} - r < \bar{t} + r < T$ . We fix  $\bar{y} \in \mathbb{R}^m$ ,  $\eta > 0$  and consider a solution  $\chi = w_\delta \in \mathcal{C}^2$  of the  $\delta$ -cell problem (23) at  $(\bar{x}, D_x \psi(\bar{t}, \bar{x}), D_{xx}^2 \psi(\bar{t}, \bar{x}))$  (see Theorem 3.1) such that

$$|\delta \chi(y) + \bar{H}(\bar{x}, D_x \psi(\bar{t}, \bar{x}), D_{xx}^2 \psi(\bar{t}, \bar{x}))| \leq \eta \quad \forall y \in B(\bar{y}, r). \quad (33)$$

We define the perturbed test function as

$$\psi^\varepsilon(t, x, y) := \psi(t, x) + \varepsilon \chi(y).$$

Observe that

$$\limsup_{\varepsilon \rightarrow 0, t' \rightarrow \bar{t}, x' \rightarrow \bar{x}, y' \rightarrow \bar{y}} V^\varepsilon(t', x', y') - \psi^\varepsilon(t', x', y') = \bar{V}(t, x) - \psi(t, x).$$

By a standard argument in viscosity solution theory (see [6, Lemma V.1.6], [20]) we get that there exist sequences  $\varepsilon_n \rightarrow 0$  and  $(t_n, x_n, y_n) \in \bar{B} := B((\bar{t}, \bar{x}, \bar{y}), r)$  such that:

$$\begin{aligned} (t_n, x_n, y_n) &\rightarrow (\bar{t}, \bar{x}, \bar{y}), \text{ for some } y \in B(\bar{y}, r), \\ V^{\varepsilon_n}(t_n, x_n, y_n) - \psi^{\varepsilon_n}(t_n, x_n, y_n) &\rightarrow \bar{V}(\bar{t}, \bar{x}) - \psi(\bar{t}, \bar{x}), \\ (t_n, x_n, y_n) &\text{ is a strict maximum of } V^{\varepsilon_n} - \psi^{\varepsilon_n} \text{ in } \bar{B}. \end{aligned}$$

Then, using the fact that  $V^\varepsilon$  is a subsolution to (9), we get

$$-\psi_t + H^{\varepsilon_n}(x_n, y_n, D_x \psi, D_{xx}^2 \psi, 0) + \lambda V^{\varepsilon_n}(t_n, x_n, y_n) - \mathcal{L}(x_n, y_n, D_y \chi, D_{yy}^2 \chi) \leq 0 \quad (34)$$

where the derivatives of  $\psi$  and  $\chi$  are computed respectively in  $(t_n, x_n)$  and in  $y_n$ . Using the fact that  $\chi$  solves the  $\delta$ -cell problem (23), we obtain

$$\begin{aligned} &-\psi_t(t_n, x_n) - \delta \chi(y_n) + \lambda V^{\varepsilon_n}(t_n, x_n, y_n) \leq \\ &\leq H(\bar{x}, y_n, D_x \psi(\bar{t}, \bar{x}), D_{xx}^2 \psi(\bar{t}, \bar{x}), 0) - H^{\varepsilon_n}(x_n, y_n, D_x \psi(t_n, x_n), D_{xx}^2 \psi(t_n, x_n), 0) \\ &+ \mathcal{L}(x_n, y_n, D_y \chi, D_{yy}^2 \chi) - \mathcal{L}(\bar{x}, y_n, D_y \chi, D_{yy}^2 \chi). \end{aligned}$$

By taking the limit as  $n \rightarrow +\infty$  the r.h.s. of this inequality cancel out. Next we use (33) to replace  $-\delta \chi$  with  $\bar{H} - \eta$  and get that the left hand side of (32) is  $\leq \eta$ . Finally, by letting  $\eta \rightarrow 0$  we obtain (32).

#### Step 4 (Behaviour of $\bar{V}$ and $\underline{V}$ at time $T$ ).

We show that, at every  $x$ ,  $\bar{V}(x, T) \leq \bar{g}(x)$  and  $\underline{V}(x, T) \geq \bar{g}(x)$ .

We start with  $\bar{V}$ . If  $g = g(x, y)$  and the coefficients  $b, \tau$  are independent of  $x$ , we reproduce and complete the argument given in [7, Thm 5.1, Step 4]. We fix  $\bar{x} \in \mathbb{R}^m$  and denote  $w_r$  the unique bounded classical solution to

$$\begin{cases} w_t - \mathcal{L}(y, Dw, D^2 w) = 0 & \text{in } (0, +\infty) \times \mathbb{R}^m \\ w^r(0, y) = \sup_{\{|x-\bar{x}| \leq r\}} g(x, y) & \text{in } \mathbb{R}^m. \end{cases} \quad (35)$$

By Proposition 3.1,  $\lim_{t \rightarrow +\infty} w_r(t, y) = \int_{\mathbb{R}^m} \sup_{|x-\bar{x}| \leq r} g(x, y) d\mu(y) =: \bar{g}_r(\bar{x})$  locally uniformly in  $y$ , where  $\mu$  is the invariant probability measure on  $\mathbb{R}^m$  for the process  $Y$ . Note that  $\sup_{|x-\bar{x}| \leq r_1} g(x, y) \geq \sup_{|x-\bar{x}| \leq r_2} g(x, y)$  for  $r_1 \geq r_2 \geq 0$ ,  $y \in \mathbb{R}^m$  and  $\lim_{r \rightarrow 0} \sup_{|x-\bar{x}| \leq r} g(x, y) = g(\bar{x}, y)$ . So, by monotone convergence theorem,  $\lim_{r \rightarrow 0} \bar{g}_r(\bar{x}) = \bar{g}(\bar{x})$ .

We fix  $r > 0$  and a constant  $K_r > 0$  such that  $V^\varepsilon(t, x, y) \leq K_r$  and  $|g(x, y)| \leq K_r/2$  for every  $\varepsilon > 0$ ,  $x \in \overline{B_r} := \overline{B(\bar{x}, r)}$ ,  $y \in \mathbb{R}^m$  and  $t \in [0, T]$ . Observe that this is possible by estimates (15) and assumption (8). Moreover we fix a smooth nonnegative function  $\psi_r$  such that  $\psi_r(\bar{x}) = 0$  and  $\psi_r(x) + \inf_y g(x, y) \geq K_r$  for every  $x \in \partial B_r$ . Let  $C_r$  be a positive constant such that

$$|H^\varepsilon(x, y, D\psi_r(x), D^2\psi_r(x), 0)| \leq C_r \quad \text{for } x \in \overline{B_r}, y \in \mathbb{R}^m \text{ and } \varepsilon > 0$$

where  $H^\varepsilon$  is defined in (12). Note that such constant exists due to assumptions (5) and (8).

We define the function

$$\psi_r^\varepsilon(t, x, y) = w_r \left( \frac{T-t}{\varepsilon}, y \right) + \psi_r(x) + C_r(T-t)$$

and we claim that it is a supersolution to the parabolic problem

$$\begin{cases} -V_t + F^\varepsilon \left( x, y, V, D_x V, \frac{D_y V}{\varepsilon}, D_{xx}^2 V, \frac{D_{yy}^2 V}{\varepsilon}, \frac{D_{xy}^2 V}{\sqrt{\varepsilon}} \right) = 0 & \text{in } (0, T) \times B_r \times \mathbb{R}^m \\ V(t, x, y) = K_r & \text{in } (0, T) \times \partial B_r \times \mathbb{R}^m \\ V(T, x, y) = g(x, y) & \text{in } \overline{B_r} \times \mathbb{R}^m \end{cases} \quad (36)$$

where  $F^\varepsilon$  is defined in (11). Indeed

$$\begin{aligned} & -(\psi_r^\varepsilon)_t + F^\varepsilon \left( x, y, D_x \psi_r^\varepsilon, \frac{D_y \psi_r^\varepsilon}{\varepsilon}, D_{xx}^2 \psi_r^\varepsilon, \frac{D_{yy}^2 \psi_r^\varepsilon}{\varepsilon}, \frac{D_{xy}^2 \psi_r^\varepsilon}{\sqrt{\varepsilon}} \right) = \\ & = \frac{1}{\varepsilon} [(w_r)_t - \mathcal{L}(y, Dw_r, D^2 w_r)] + C_r + H^\varepsilon(y, x, D\psi_r(x), D^2\psi_r(x), 0) \geq 0. \end{aligned}$$

Moreover  $\psi_r^\varepsilon(T, x, y) = \sup_{|x-\bar{x}| \leq r} g(x, y) + \psi_r(x) \geq g(x, y)$ . Finally, observe that the constant function  $\inf_y \sup_{|x-\bar{x}| \leq r} g(x, y)$  is always a subsolution to (35) and then by a standard comparison principle we obtain  $w_r(t, y) \geq \inf_y \sup_{|x-\bar{x}| \leq r} g(x, y)$ . This implies

$$\psi_r^\varepsilon(t, x, y) \geq \inf_y \sup_{|x-\bar{x}| \leq r} g(x, y) + K_r - \inf_y g(x, y) + C_r(T-t) \geq K_r \quad \forall x \in \partial B_r.$$

Then  $\psi^\varepsilon$  is a supersolution to (36). For our choice of  $K_r$ , we get that  $V^\varepsilon$  is a subsolution to (36). Moreover both  $V^\varepsilon$  and  $\psi_r^\varepsilon$  are bounded in  $[0, T] \times \overline{B_r} \times \mathbb{R}^m$ , because of the estimate (15), of the boundedness of  $w_r$  and of the regularity of  $\psi_r$ . So, a standard comparison principle for viscosity solutions gives

$$V^\varepsilon(t, x, y) \leq \psi_r^\varepsilon(t, x, y) = w_r \left( \frac{T-t}{\varepsilon}, y \right) + \psi_r(x) + C_r(T-t)$$

for every  $r > 0$ ,  $\varepsilon > 0$ ,  $(t, x, y) \in ([0, T] \times \overline{B_r} \times \mathbb{R}^m)$ . We compute the upper limit of both sides of the previous inequality as  $(\varepsilon, t, x, y) \rightarrow (0, t', x', y')$  for  $t' \in (0, T)$ ,  $x' \in B_r$ ,  $y' \in \mathbb{R}^m$  and get

$$\overline{V}(t', x') \leq \overline{g}_r(\bar{x}) + \psi_r(x') + C_r(T-t').$$

Then, taking the upper limit for  $(t', x') \rightarrow (T, \bar{x})$ , we obtain  $\overline{V}(T, \bar{x}) \leq \overline{g}_r(\bar{x})$ , for every  $r > 0$ . This permits to conclude, sending  $r \rightarrow 0$ .

If  $g$  does not depend on  $y$ , we get that the function  $\psi_r^\varepsilon(t, x, y) = \sup_{|x-\bar{x}| \leq r} g(x) + \psi_r(x) + C_r(T-t)$  is a supersolution to (36), and the conclusion easily follows.



The proof for  $\underline{V}$  is completely analogous, once we replace the Cauchy problem (35) with

$$\begin{cases} w_t - \mathcal{L}(y, Dw, D^2w) = 0 & \text{in } (0, +\infty) \times \mathbb{R}^m \\ w(0, y) = \inf_{\{|x-\bar{x}|\leq r\}} g(x, y) & \text{in } \mathbb{R}^m. \end{cases}$$

**Step 5 (Uniform convergence).**

Observe that both  $\underline{V}$  and  $\bar{V}$  satisfy the same quadratic growth condition (30). Moreover we assume that the Comparison Principle holds for the problem (21), so  $\underline{V} \geq \bar{V}$ . Therefore, since by definition  $\underline{V} \leq \bar{V}$ , we conclude  $\underline{V} = \bar{V} =: V$ . This implies that  $V$  is continuous and that, by the definition of semilimits,  $V^\varepsilon$  converges loc. uniformly to  $V$  (see [6, Lemma V.1.9]).  $\square$

## 4 Examples and applications

In this section we look for a control problem whose value function is the limit  $V$  of the value functions  $V^\varepsilon$ . It is enough to represent the effective Hamiltonian  $\bar{H}$  as a HJB Hamiltonian, that is,

$$\bar{H}(x, p, P) = \min_{\tilde{u} \in \tilde{U}} \left\{ -\text{trace}(\bar{\sigma}\bar{\sigma}^T(x, \tilde{u})P) - \bar{f}(x, \tilde{u}) \cdot p - \bar{l}(x, \tilde{u}) \right\}, \quad (37)$$

for some compact set  $\tilde{U}$  and continuous  $\bar{f}, \bar{\sigma}, \bar{l}$ , with  $\bar{f}, \bar{\sigma}$  Lipschitz in  $x$ . Indeed in this case the uniqueness of the viscosity solution to the Cauchy problem for HJB equations gives the formula

$$V(t, x) := \sup_{\tilde{u}} \mathbf{E} \left[ \int_t^T \bar{l}(X_s, \tilde{u}_s) ds + \langle g \rangle(X_T) \right],$$

for  $X_s$  solving

$$dX_s = \bar{f}(X_s, \tilde{u}_s) ds + \sqrt{2}\bar{\sigma}(X_s, \tilde{u}_s) dW_s, \quad X_t = x,$$

where we used the notation

$$\langle \phi \rangle(x) := \int_{\mathbb{R}^m} \phi(x, y) d\mu_x(y), \quad \mu_x \text{ invariant measure of the fast subsystem.}$$

Therefore we get an *effective control problem* which is the variational limit of the two-scale optimal control problem. If  $\sigma \equiv 0$  the Hamiltonian is of first order,  $\bar{H} = \bar{H}(x, p)$ , and so the effective problem is deterministic, namely,

$$\sup_{\tilde{u}} \left( \int_t^T \bar{l}(X_s, \tilde{u}_s) ds + \langle g \rangle(X_T) \right), \quad \dot{X}_s = \bar{f}(X_s, \tilde{u}_s).$$

Note that a HJB representation of the effective Hamiltonian always exists because  $\bar{H}(x, p, X)$  is concave with respect to  $p$  and  $X$ , by convex duality. However, such an abstract representation of the effective problem is not very useful and we rather seek an explicit one given by averaging the data with respect to the invariant measure  $\mu_x$ . In the rest of the section we present various cases where explicit formulas for the effective control problem can be given. The first leads to the most natural linear averaging of the data, as for the uncontrolled systems. The others, instead, lead to different kinds of averages.

#### 4.1 Controls decoupled from the fast variables

In this section we consider slow systems in split form

$$dX_s = [f_1(X_s, u_s) + f_2(X_s, Y_s)] ds + \sqrt{2} [\sigma_1(X_s, u_s) + \sigma_2(X_s, Y_s)] dW_s, \quad (38)$$

and cost functionals with split running cost

$$J^\varepsilon(t, x, y, u) := \mathbf{E} \left[ \int_t^T (l_1(X_s, u_s) + l_2(X_s, Y_s)) ds + g(X_T, Y_T) \mid X_t = x, Y_t = y \right]. \quad (39)$$

In other words, in all the data the control  $u_s$  is decoupled from the fast variables  $Y_s$ . We also assume that

$$\sigma_1(x, u)\sigma_2^T(x, y) = 0 \quad \text{for all } x, u, y. \quad (40)$$

This is a condition of uncorrelation of the two diffusion terms  $\sigma_1 dW_s$  and  $\sigma_2 dW_s$ . In fact it is satisfied if the diffusion term  $\sigma dW_s$  of the slow system is of the form

$$\tilde{\sigma}_1(X_s, u_s) dW_s^1 + \tilde{\sigma}_2(X_s, Y_s) dW_s^2 \quad \text{with } W^1 \text{ and } W^2 \text{ independent.}$$

Indeed, in this case

$$\sigma(x, y, u) = \begin{pmatrix} \tilde{\sigma}_1(x, u) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\sigma}_2(x, y) \end{pmatrix}.$$

By the decoupling assumption on the data and the uncorrelation condition (40), the effective Hamiltonian is

$$\begin{aligned} \bar{H}(x, p, P) &:= \int H(x, y, p, P, 0) d\mu_x(y) \\ &= \min_{u \in U} \{ -\text{trace}[P\sigma_1\sigma_1^T(x, u)] - p \cdot f_1(x, u) - l_1(x, u) \} \\ &\quad - \text{trace} \left[ P \int \sigma_2\sigma_2^T(x, y) d\mu_x(y) \right] - p \cdot \int f_2(x, y) d\mu_x(y) - \int l_2(x, y) d\mu_x(y). \end{aligned}$$

The last condition we assume is the existence of a Lipschitz square root of the matrix  $\langle \sigma_2\sigma_2^T \rangle$ , i.e., a Lipschitz continuous  $n \times r$  matrix valued function  $\bar{\sigma}_2(x)$  such that

$$\bar{\sigma}_2(x)\bar{\sigma}_2(x)^T = \int_{\mathbb{R}^m} \sigma_2(x, y)\sigma_2^T(x, y) d\mu_x(y).$$

Then we have the representation (37) for  $\bar{H}$  with  $\tilde{U} = U$  and

$$\bar{\sigma} = \sigma_1 + \bar{\sigma}_2 = \sigma_1 + \langle \sigma_2\sigma_2^T \rangle^{1/2}, \quad \bar{f} = f_1 + \langle f_2 \rangle = \langle f \rangle, \quad \bar{l} = l_1 + \langle l_2 \rangle = \langle l \rangle.$$

We therefore get the effective optimal control problem governed by the system

$$dX_s = [f_1(X_s, u_s) + \langle f_2 \rangle(X_s)] ds + \sqrt{2} [\sigma_1(X_s, u_s) + \bar{\sigma}_2(X_s)] dW_s, \quad (41)$$

and with payoff functional

$$\bar{J}(t, x, u) := \mathbf{E} \left[ \int_t^T (l_1(X_s, u_s) + \langle l_2 \rangle(X_s)) ds + \langle g \rangle(X_T) \mid X_t = x \right]. \quad (42)$$

The convergence of the value functions  $V^\varepsilon$  to the value function  $V$  of this limit control problem was proved by Kushner [35] in the special case of uncontrolled diffusion matrix, that is,  $\sigma_1 \equiv 0$ , and under more stringent assumptions than ours on  $g$  and on the fast subsystem (17). Note that in this case  $\overline{\sigma\sigma^T} = \langle \sigma\sigma^T \rangle$ , so all terms of the effective problem are the linear average of the corresponding data of the two-scale problem. Kushner used probabilistic methods completely different from ours. We recover his result under the additional condition that  $f$  and  $l$  are bounded, by Theorem 3.2 and point (v) of Proposition 3.2.

**Remark 4.1.** If the initial system is deterministic for frozen  $Y_t$ , that is,  $\sigma = \sigma_1 + \sigma_2 \equiv 0$ , then the effective control problem is deterministic. The effective system is

$$\dot{X}_s = f_1(X_s, u_s) + \langle f_2 \rangle(X_s), \quad X_t = x,$$

and the effective payoff functional is

$$\bar{J}(t, x, u) := \int_t^T [l_1(X_s, u_s) + \langle l_2 \rangle(X_s)] ds + \langle g \rangle(X_T).$$

## 4.2 Optimal economic growth

The most classical model of optimal growth of a one-sector economy goes back to Ramsey, see [30, 17, 41]. Let  $K_s$  denote the capital per worker and  $u_s$  the consumption per worker. The dynamics is

$$\dot{K}_s = f(K_s) - u_s, \quad \text{for } s > t, \quad K_t = k.$$

The running cost depends only on the control,  $l = l(u)$ , with the utility function  $l$  strictly increasing and concave, and we take a Hyperbolic Absolute Risk Aversion (HARA) function

$$l(u) = \theta \frac{u^\gamma}{\gamma}, \quad \theta > 0, \quad 0 < \gamma < 1. \quad (43)$$

The payoff functional is

$$J(t, k, u) = \int_t^T e^{-\lambda(T-s)} \theta \frac{u_s^\gamma}{\gamma} ds + e^{-\lambda(T-t)} g(K_T), \quad \lambda \geq 0.$$

The constraint on the control depends in general on the state

$$0 \leq u_s \leq R(K_s), \quad R > 0,$$

and often it is taken  $R(K_s) = f(K_s) + \delta K_s$ , where  $\delta$  is the rate of depreciation of the capital. The Hamiltonian is

$$H(k, p) = \min_{0 \leq u \leq R(k)} \left\{ up - \theta \frac{u^\gamma}{\gamma} \right\} - f(k)p.$$

The min is always attained in the interior of  $U = [0, R(k)]$  or at the right endpoint, and the two cases give

$$H(k, p) = \begin{cases} \theta^{\frac{1}{1-\gamma}} \frac{\gamma-1}{\gamma} p^{\frac{\gamma}{\gamma-1}} - f(k)p, & p \geq \theta R(k)^{\gamma-1}, \\ R(k)p - \theta \frac{R(k)^\gamma}{\gamma} - f(k)p, & p < \theta R(k)^{\gamma-1}. \end{cases}$$

**The two-scale problem and its limit.** Now we consider the problem with random parameters

$$\begin{cases} \dot{K}_s = f(K_s, Y_s) - u_s, & K_t = k, \\ dY_s = \frac{1}{\varepsilon} b(Y_s) ds + \sqrt{\frac{2}{\varepsilon}} \tau(Y_s) dW_s & Y_t = y, \end{cases}$$

$$J^\varepsilon(t, k, y, u_\cdot) = \mathbf{E} \left[ \int_t^T e^{-\lambda(T-s)} \theta(Y_s) \frac{u_s^\gamma}{\gamma} ds + e^{-\lambda(T-t)} g(K_T, Y_T) \right],$$

assuming for simplicity that the dynamics of  $Y_s$  does not depend on the capital. Then  $\mu$  is independent of  $k$  and we denote  $\langle \phi \rangle(k) := \int_{\mathbb{R}^m} \phi(k, y) d\mu(y)$ . To fit in the assumptions of the previous sections we assume  $R = \text{constant}$ , although the theory could be extended to cover the case of  $R(\cdot)$  Lipschitz and bounded. By averaging  $H(k, y, p)$  with respect to  $\mu(y)$  we get the effective Hamiltonian

$$\bar{H}(k, p) = \bar{h}(p) - \langle f \rangle(k)p,$$

where

$$\bar{h}(p) := \frac{\gamma-1}{\gamma} p^{\frac{\gamma}{\gamma-1}} \int_{\{y: \theta(y) \leq pR^{1-\gamma}\}} \theta(y)^{\frac{1}{1-\gamma}} d\mu(y) + \int_{\{y: \theta(y) > pR^{1-\gamma}\}} \left[ Rp - \theta(y) \frac{R^\gamma}{\gamma} \right] d\mu(y).$$

This corresponds to an effective control problem with linearly averaged dynamics

$$\dot{K}_s = \langle f \rangle(K_s) - u_s, \quad u_s \geq 0,$$

and the effective payoff

$$\bar{J}(t, k, u_\cdot) = \int_t^T e^{-\lambda(T-s)} \bar{l}(u_s) ds + e^{-\lambda(T-t)} \langle g \rangle(K_T),$$

where the effective utility function  $\bar{l}$  is the Legendre-Fenchel conjugate of  $\bar{h}$ , i.e.,

$$\bar{l}(u) := \inf_{p \geq 0} \{ up - \bar{h}(p) \}.$$

Note that, for any  $p$  "not too small", in the sense that  $\mu\{y : \theta(y) \leq pR^{1-\gamma}\} = 1$ , the effective Hamiltonian takes the much simpler form

$$\bar{H}(k, p) = \langle \theta^{\frac{1}{1-\gamma}} \rangle \frac{\gamma-1}{\gamma} p^{\frac{\gamma}{\gamma-1}} - \langle f \rangle(k)p,$$

which is the same as in the constant coefficients case, but the effective parameter  $\bar{\theta}$  is the generalized harmonic *nonlinear power-like average*

$$\bar{\theta} = \langle \theta^{\frac{1}{1-\gamma}} \rangle^{1-\gamma},$$

different from the preceding Section 4.1.

**A remark on the limit of the optimal control.** The behavior of the optimal controls of the problem with random parameters as  $\varepsilon \rightarrow 0$  is much harder to understand than the convergence of the Hamiltonians and value function. To illustrate what can be expected we propose some formal calculations on a special case of the economic growth model.

We consider  $l$  as in (43) with  $\gamma = 1/2$  and assume that also the terminal cost is a HARA function of  $k$ , whereas  $f$  is linear in  $k$

$$g(k, y) = a(y)2k^{1/2}, \quad f(k, y) = r(y)k, \quad a(y) \geq a_o > 0.$$

We allow all locally integrable controls  $u_s \geq 0$  (i.e.,  $R = +\infty$ ), and we impose the natural state constraint

$$K_s \geq 0 \quad \text{for all } s.$$

This is equivalent to stopping the process when the capital reaches 0 and getting a null terminal cost. In this case we can exhibit an explicit solution of the effective Cauchy problem. However the model problem does not fit into the assumptions of the previous sections for two reasons. One is that the HJB equation holds only for  $k > 0$  and there is a boundary condition at  $k = 0$  (either of state-constrained or Dirichlet, see [6]). This requires an additional argument in the convergence analysis, see [7] for a different example. The second difficulty is the unboundedness of the control set, leading to different growth conditions on the Hamiltonian and the value functions. We believe the convergence result is still true under suitable coercivity assumptions on the costs, but the proof would require some nontrivial variants to the one presented here.

The HJB equation for  $V^\varepsilon$  under the current assumptions is

$$-V_t^\varepsilon + \min_{u \geq 0} \{uV_k^\varepsilon - 2\theta(y)\sqrt{u}\} - r(y)kV_k^\varepsilon + \lambda V = \frac{1}{\varepsilon} \mathcal{L}(y, D_y V^\varepsilon, D_{yy}^2 V^\varepsilon)$$

and the nonlinear term in the Hamiltonian is

$$\min_{u \geq 0} \{up - 2\theta(y)\sqrt{u}\} = -\frac{\theta(y)^2}{p},$$

the min being attained at  $(p/\theta)^{-2}$ . The effective HJB equation suggested by the previous sections is

$$-V_t - \frac{\langle \theta^2 \rangle}{V_k} - \langle r \rangle k V_k + \lambda V = 0$$

with the terminal condition  $V(T, k) = \langle a \rangle 2k^{1/2}$ . It is natural to look for a solution of the HARA form  $V(t, k) = 2k^{1/2}v(t)$ . Then we get the ODE

$$-2\dot{v} - \frac{\langle \theta^2 \rangle}{v} - \langle r \rangle v + 2\lambda v = 0.$$

The sign of the singular term ensures the existence in  $(-\infty, T]$  of a solution satisfying  $v(T) = \langle a \rangle > 0$ . Moreover, the control where the min is attained in the Hamiltonian is

$$u^*(t, k) = \left( \frac{V_k(t, k)}{\theta} \right)^{-2} = \frac{\langle \theta^2 \rangle}{v(t)^2} k,$$

and this is therefore the optimal (feedback) control of the effective problem. Next we guess that also  $V^\varepsilon$  has the HARA form  $V^\varepsilon(t, k, y) = 2k^{1/2}v^\varepsilon(t, y)$ . If it is so, then the optimal feedback control for the problem with random parameters is

$$u_\varepsilon^*(t, k, y) = \left( \frac{V_k^\varepsilon(t, k, y)}{\theta(y)} \right)^{-2} = \frac{\theta(y)^2}{v^\varepsilon(t, y)^2} k.$$

Motivated by the previous sections we conjecture that  $V^\varepsilon \rightarrow V$  (say, pointwise) and therefore

$$\lim_{\varepsilon \rightarrow 0} u_\varepsilon^*(t, k, y) = \frac{\theta(y)^2}{v(t)^2} k.$$

Note that this limit is different from the optimal control  $u^*$  of the limit problem (if  $\theta$  is not constant...). However,  $u^*$  is the average of the limit of the  $u_\varepsilon^*$  as well as the limit of the averages:

$$u^*(t, k) = \int_{\mathbb{R}^m} \lim_{\varepsilon \rightarrow 0} u_\varepsilon^*(t, k, y) d\mu(y) = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} u_\varepsilon^*(t, k, y) d\mu(y).$$

A similar result was proved in [7] (see also [26]) for the Merton problem of optimizing the investment in a portfolio containing a riskless bond and a risky stock with fast stochastic volatility.

### 4.3 An advertising model

The most classical model of optimal advertising is due to Vidale and Wolfe. In a market with total sales  $M > 0$  the sales of a given firm are  $S_s$  and  $0 \leq u_s \leq R$  is the advertising effort. The dynamics is

$$\dot{S}_s = \gamma(M - S_s)u_s - \delta S_s, \quad \text{for } s > t, \quad S_t = S$$

where  $\gamma > 0$  measures the effectiveness of the advertising (which acts on the unsold portion of the market) and  $\delta > 0$  measures the loss due to forgetting (which acts on the sold portion of the market). Note that the set  $[0, M]$  is invariant for this control system.

The firm wants to maximize its total revenue net of advertising costs. We assume the income proportional to the sales and the advertising cost quadratic in  $u$ , and so the payoff functional is

$$J(t, S, u) = \int_t^T (rS_s - \theta u_s^2) ds,$$

with  $\theta, r > 0$ , see [41, 33]. The Hamiltonian is

$$H(S, p) = \min_{0 \leq u \leq R} \{ \theta u^2 - \gamma(M - S)up \} + \delta Sp - rS.$$

We expect that the value function  $V(t, S) := \sup_u J(t, S, u)$  be nondecreasing with respect to the initial sales  $S$ . Then  $V_S \geq 0$ , and for  $p \geq 0$  we compute

$$H(S, p) = \begin{cases} -\frac{\gamma^2}{4\theta}(M - S)^2 p^2 + \delta Sp - rS, & 0 \leq p \leq \frac{2\theta R}{\gamma(M - S)}, \\ \theta R^2 - \gamma(M - S)Rp + \delta Sp - rS, & p > \frac{2\theta R}{\gamma(M - S)}. \end{cases}$$

**The two-scale problem and the effective Hamiltonian.** Now we consider the problem with random parameters

$$\begin{cases} \dot{S}_s = \gamma(Y_s)(M(Y_s) - S_s)u_s - \delta(Y_s)S_s, & S_t = S, \\ dY_s = \frac{1}{\varepsilon} b(Y_s) ds + \sqrt{\frac{2}{\varepsilon}} \tau(Y_s) dW_s & Y_t = y, \end{cases}$$

$$J^\varepsilon(t, S, y, u) = \mathbf{E} \left[ \int_t^T (r(Y_s)S_s - \theta(Y_s)u_s^2) ds \right],$$

assuming for simplicity that the dynamics of  $Y_s$  does not depend on the sales. Then  $\mu$  is independent of  $S$  and we denote  $\langle \phi \rangle := \int_{\mathbb{R}^m} \phi(S, y) d\mu(y)$ . By averaging  $H(S, y, p)$  with respect to  $\mu(y)$  we get the following effective Hamiltonian for the limit as  $\varepsilon \rightarrow 0$ :

$$\bar{H}(S, p) = \bar{h}(S, p) + \langle \delta \rangle Sp - \langle r \rangle S, \quad (44)$$

where

$$\begin{aligned} \bar{h}(S, p) := & -p^2 \int_{\{y: p\gamma(M-S) \leq 2\theta R\}} \frac{\gamma(y)^2}{4\theta(y)} (M(y) - S)^2 d\mu(y) + \\ & \int_{\{y: p\gamma(M-S) > 2\theta R\}} [\theta(y)R^2 - \gamma(y)(M(y) - S)Rp] d\mu(y). \end{aligned} \quad (45)$$

Note that, for any  $p$  "not too large", in the sense that  $\mu\{y : p\gamma(y)(M(y) - S) \leq 2\theta(y)R\} = 1$ , the effective Hamiltonian takes the much simpler form

$$\bar{H}(S, p) = -\langle \frac{\gamma^2}{4\theta} (M - S)^2 \rangle p^2 + \langle \delta \rangle Sp - \langle r \rangle S. \quad (46)$$

In some cases this form is the same as in the constant coefficients case, with suitable effective parameters  $\bar{\theta}$  and  $\bar{\gamma}$ . For instance, if  $\gamma$  and  $M$  are constants independent of  $Y_s$  the first term is

$$\bar{h}(S, p) = -\frac{\gamma^2}{4\theta} (M - S)^2 p^2 \quad \text{for } \bar{\theta} = \langle \frac{1}{\theta} \rangle^{-1},$$

i.e., the effective coefficient  $\bar{\theta}$  is the *harmonic average* of  $\theta(y)$ , instead of the linear averages that appeared in Section 4.1. Similarly, if  $M$  and  $\theta$  do not depend on the random parameters  $Y_s$  the first term is

$$\bar{h}(S, p) = -\frac{\bar{\gamma}^2}{4\theta} (M - S)^2 p^2 \quad \text{for } \bar{\gamma} = \sqrt{\langle \gamma^2 \rangle},$$

i.e., the effective  $\bar{\gamma}$  is a *nonlinear power-like average* of  $\gamma(y)$ .

**The effective control problem and its interpretation.** We can find an explicit effective control problem letting the upper bound  $R$  of the control go to  $+\infty$  in the case of  $M$  constant. Let  $\bar{H}^R$  denote the Hamiltonian given by (44) and (45) and  $\bar{H}^\infty$  the one given by (46) for all  $S \in [0, M], p \geq 0$ . Let  $V^R$  be the viscosity solution of the effective Cauchy problem

$$-V_t^R + \bar{H}^R(S, V_S^R) = 0 \text{ in } (0, T) \times (0, M), \quad V^R(T, S) = 0,$$

whose uniqueness, despite the lack of boundary conditions at  $S = 0, M$ , follows from the invariance of the interval  $[0, M]$  (see Proposition 3.1 in [7]). Now define

$$V^\infty(t, S) := \sup_{u \in L^\infty([0, T], [0, +\infty))} \int_t^T (\langle r \rangle S_s - u_s^2) ds, \quad (47)$$

where  $S_s$  is the state of the system

$$\dot{S}_s = \langle \frac{\gamma^2}{\theta} \rangle^{\frac{1}{2}} (M - S_s) u_s - \langle \delta \rangle S_s \text{ for } s > t, \quad S_t = S. \quad (48)$$

Then, by standard methods as in [8],  $V^\infty$  is Lipschitz in  $S$  and  $\frac{1}{2}$ -Hölder continuous in  $t$  and it is the unique viscosity solution Lipschitz in  $S$  of the Cauchy problem

$$-V_t^\infty + \overline{H}^\infty(S, V_S^\infty) = 0 \text{ in } (0, T) \times (0, M), \quad V^\infty(T, S) = 0, \quad (49)$$

as well as the minimal supersolution. We claim that  $V^\infty = \sup_{R>0} V^R = \lim_{R \rightarrow \infty} V^R$ . In fact,  $\overline{H}^R \geq \overline{H}^\infty$  and the Comparison Principle imply  $\sup_{R>0} V^R \leq V^\infty$ . On the other hand,  $\overline{H}^R$  converges locally uniformly to  $\overline{H}^\infty$ . Then a standard argument in the theory of viscosity solutions (see, e.g., Chapter V of [6]) implies that  $\sup_{R>0} V^R$  is a supersolution of (49), therefore it is  $\geq V^\infty$ , which completes the proof of the claim.

Observe that the limit control problem (47) (48) with value function  $V^\infty$  is a Vidale - Wolfe problem with constant (effective) parameters, although some effective coefficients are obtained by a nonlinear average, as remarked above. In particular, if  $\theta$  is also constant, the effective control problem can be written as

$$\dot{S}_s = \sqrt{\langle \gamma^2 \rangle} (M - S_s) u_s - \langle \delta \rangle S_s, \quad \overline{J}(t, S, u.) = \int_t^T (\langle r \rangle S_s - \theta u_s^2) ds,$$

whereas for  $\gamma$  constant it becomes

$$\dot{S}_s = \gamma (M - S_s) u_s - \langle \delta \rangle S_s, \quad \overline{J}(t, S, u.) = \int_t^T \left( \langle r \rangle S_s - \langle \frac{1}{\theta} \rangle^{-1} u_s^2 \right) ds.$$

Finally we note that, if  $\gamma$  and  $\theta$  are constant but  $M$  does depend on  $Y_s$ , the Hamiltonian  $\overline{H}^\infty$  in (46) can be formally associated to the following control problem

$$\dot{S}_s = \gamma \sqrt{\langle M^2 \rangle - 2\langle M \rangle S_s + S_s^2} u_s - \langle \delta \rangle S_s, \quad \overline{J}(t, S, u.) = \int_t^T (\langle r \rangle S_s - \theta u_s^2) ds.$$

which is not a Vidale - Wolfe problem.

## References

- [1] O. Alvarez, M. Bardi: *Viscosity solutions methods for singular perturbations in deterministic and stochastic control*, SIAM J. Control Optim. 40 (2001/02), 1159–1188.
- [2] O. Alvarez, M. Bardi: *Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result*, Arch. Ration. Mech. Anal. 170 (2003), 17–61.
- [3] O. Alvarez, M. Bardi: *Ergodicity, stabilization, and singular perturbations for Bellman-Isaacs equations*, Mem. Amer. Math. Soc., 204 (2010), no. 960.
- [4] O. Alvarez, M. Bardi, C. Marchi: *Multiscale problems and homogenization for second-order Hamilton-Jacobi equations*, J. Differential Equations 243 (2007), 349–387.
- [5] O. Alvarez, M. Bardi, C. Marchi: *Multiscale singular perturbations and homogenization of optimal control problems*, in Geometric control and nonsmooth analysis (ed. F. Ancona et al.), Series on Advances in Mathematics for Applied Sciences, vol. 76 (World Scientific, 2008).



- [6] M. Bardi, I. Capuzzo-Dolcetta: **Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations**, Birkäuser, Boston, 1997.
- [7] M. Bardi, A. Cesaroni, L. Manca: *Convergence by viscosity methods in multiscale financial models with stochastic volatility*, SIAM J. Finan. Math. 1 (2010), pp. 230-265.
- [8] M. Bardi, F. Da Lio: *On the Bellman equation for some unbounded control problems*, NoDEA Nonlinear Differential Equations Appl. 4 (1997), no. 4, 491–510.
- [9] G. Barles, S. Biton, M. Bourgoing, O. Ley: *Uniqueness results for quasilinear parabolic equations through viscosity solutions' methods*, Calc. Var. Partial Differential Equations 18 (2003), no. 2, 159–179.
- [10] A. Bensoussan: **Perturbation methods in optimal control**, John Wiley & Sons, Montrouge, 1988.
- [11] A. Bensoussan, G. Blankenship: *Controlled diffusions in a random medium*, Stochastics 24 (1988), no. 2, 87–120.
- [12] J.F. Bonnans, A. Shapiro: **Perturbation analysis of optimization problems**, Springer Series in Operations Research. Springer-Verlag, New York, 2000
- [13] V.S. Borkar, V. Gaitsgory: *Averaging of singularly perturbed controlled stochastic differential equations*, Appl. Math. Optim. 56 (2007), 169–209.
- [14] V.S. Borkar, V. Gaitsgory: *Singular perturbations in ergodic control of diffusions*, SIAM J. Control Optim. 46 (2007), 1562–1577.
- [15] B. Bouchard, N. Touzi: *Weak Dynamic Programming Principle for Viscosity Solutions*, preprint 2009, available at HAL preprint server.
- [16] L.A. Caffarelli, P.E. Souganidis, L. Wang : *Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media*, Comm. Pure Appl. Math. 58 (2005), no. 3, 319–361.
- [17] L. Cesari: **Optimization—theory and applications**. Springer-Verlag, New York, 1983.
- [18] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, P.R. Wolenski: **Nonsmooth analysis and control theory**. Springer-Verlag, New York, 1998.
- [19] F. Comets, T. Meyre: **Calcul stochastique et modeles de diffusions**, Dunod, Paris, 2006.
- [20] M.G. Crandall, H. Ishii, P.-L. Lions: *Users guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) 27 (1992), 1–67.
- [21] F. Da Lio, O. Ley: *Uniqueness results for second-order Bellman-Isaacs equations under quadratic growth assumptions and applications*, SIAM J. Control Optim. 45 (2006), 74–106.
- [22] M.G. Dmitriev, G.A. Kurina: *Singular perturbations in control problems*. Autom. Remote Control 67 (2006), no. 1, 1–43

- [23] A.L. Dontchev, T. Zolezzi: **Well-posed optimization problems**. Lecture Notes in Mathematics, 1543. Springer-Verlag, Berlin, 1993
- [24] L. C. Evans: *The perturbed test function method for viscosity solutions of nonlinear PDE*, Proc. Roy. Soc. Edinburgh Sect. A 111 (1989), 359–375.
- [25] W. H. Fleming, H. M. Soner: **Controlled Markov processes and viscosity solutions**, 2nd edition, Springer, New York, 2006.
- [26] J.-P. Fouque, G. Papanicolaou, R. Sircar: **Derivatives in financial markets with stochastic volatility**, Cambridge University Press, Cambridge, 2000.
- [27] J.-P. Fouque, G. Papanicolaou, R. Sircar, K. Solna: *Singular perturbations in option pricing*, SIAM J. Appl. Math. 63 (2003), no. 5, 1648–1665.
- [28] J.-P. Fouque, G. Papanicolaou, R. Sircar, K. Solna: *Multiscale stochastic volatility asymptotics*, Multiscale Model. Simul. 2 (2003), no. 1, 22–42.
- [29] R.Z. Hasminskii: **Stochastic stability of differential equations**, Sijthoff & Noordhoff, Alphen aan den Rijn–Germantown, 1980.
- [30] M.D. Intriligator: **Mathematical optimization and economic theory**. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.
- [31] H. Ishii, *Uniqueness of unbounded viscosity solution of Hamilton Jacobi equations*, Indiana Univ. Math. J. 33 (1984), no. 5, 721–748.
- [32] R. Jensen and P.-L. Lions, *Some asymptotic problems in fully nonlinear elliptic equations and stochastic control*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 11 (1984), 129–176.
- [33] S. Jorgensen, G. Zaccour: **Differential Games in Marketing**, Kluwer Academic Publishers, 2004.
- [34] Y. Kabanov and S. Pergamenschikov: **Two-scale stochastic systems. Asymptotic analysis and control**, Springer-Verlag, Berlin, 2003.
- [35] H. J. Kushner: **Weak convergence methods and singularly perturbed stochastic control and filtering problems**. Birkhäuser Boston, Boston, 1990.
- [36] P.-L. Lions, M. Musiela: *Ergodicity of diffusion processes*, preprint 2002.
- [37] P.-L. Lions, G. Papanicolaou, S.R.S. Varadhan: *Homogenization of Hamilton-Jacobi equations*, Unpublished, 1986.
- [38] D.S. Naidu, *Singular perturbations and time scales in control theory and applications: an overview*, Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms 9 (2002), 233–278.
- [39] E. Pardoux, A.Yu. Veretennikov: *On the Poisson equation and diffusion approximation, I*, Ann. Probab. 29 (2001), 1061–1085.
- [40] E. Pardoux, A.Yu. Veretennikov: *On the Poisson equation and diffusion approximation, II*, Ann. Probab. 31 (2003), 1166–119.

- [41] S.P. Sethi, G.L. Thompson: **Optimal control theory. Applications to management science and economics**. Second edition. Kluwer Academic Publishers, Boston, MA, 2000
- [42] R. Sircar, T. Zariphopoulou: *Bounds and asymptotic approximations for utility prices when volatility is random*, SIAM J. Control Optim. 43 (2004/05), no. 4, 1328–1353.
- [43] R. Vinter: **Optimal control**. Birkhäuser Boston, Boston, MA, 2000.