Pure and Applied Functional Analysis
Volume 2, Number 3, 2017, 441–460

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

AN ERGODIC STACKELBERG TEAM PROBLEM FOR CONTROLLED DIFFUSIONS

VIVEK S. BORKAR

Abstract. We revisit the problem of ergodic control of singularly perturbed diffusions studied in [4] and extend the results to a Stackelberg team problem wherein a leader-follower team of controllers seeks to optimize a common ergodic cost, with the follower operating on a fast time scale with full observations and the leader on a slow time scale based only on her own observations. The main result is the validation of an averaged control system as a legitimate approximation in the limiting case when the time scale separation diverges.

1. Introduction

Singular perturbations in controlled diffusions induce two time scales, fast and slow, depending on a parameter ϵ. Intuitively, the fast components see the slow components as quasi-static and the slow components see the fast ones as quasi-equilibrated. That is, the fast dynamics is well approximated on its time scale by treating the slow variables as constant and the slow dynamics is well approximated on its time scale by averaging out its coefficients w.r.t. the asymptotic behavior of the fast components as reflected in their stationary distribution. Here 'well approximated' stands for 'within an approximation error that vanishes in the $\epsilon \downarrow 0$ limit'. This in particular allows us to use the latter averaged system arising in the $\epsilon \downarrow 0$ limit as a legitimate approximation for the singularly perturbed dynamics. This has been the basis of many works on such systems, see, e.g., the monographs [7], [8], [9]. The ergodic or long run average cost criterion has received relatively less attention in this framework, see, however, [4], [12]. This work is a sequel to [4]. It considers a 'Stackelberg team problem' situation wherein there are two controllers or 'players' distinguished from each other by deeming one a leader operating on the slow time scale and choosing control based on only her own observations, and a follower, who chooses his own control by observing the dynamics and control of both.

The paper is organized as follows. In the next section we set up the problem framework and recapitulate the key results from [4] that form the backdrop for the present contribution. Section 3 formulates the Stackelberg team problem and states and proves the main results using, among other things, a nonlinear filter that

2010 *Mathematics Subject Classification.* 91A25, 91A05, 91A15, 93E20, 60H30 .
Key words and phrases. Controlled diffusions, singular perturbations, averaging; ergodic control, Stackelberg problem, nonlinear filter.

estimates the fast component given the slow state-control processes. An Appendix sketches the derivation of the nonlinear filter in this context.

2. Background

We first summarize the results of [4] as they provide the background for the present work. In [4], we consider a coupled pair of controlled diffusions $(z^\epsilon(\cdot), x^\epsilon(\cdot))$ in $\mathcal{R}^d \times \mathcal{R}^s$, given by

$$dz^\epsilon(t) = h(z^\epsilon(t), x^\epsilon(t), u(t))dt + \gamma(z^\epsilon(t))dB(t), \tag{2.1}$$

$$dx^\epsilon(t) = \frac{1}{\epsilon} m(z^\epsilon(t), x^\epsilon(t), u(t))dt + \frac{1}{\sqrt{\epsilon}} \sigma(z^\epsilon(t), x^\epsilon(t))dW(t). \tag{2.2}$$

Here,

- for a prescribed compact metric action space A, $h : \mathcal{R}^d \times \mathcal{R}^s \times A \mapsto \mathcal{R}^d$, $\gamma : \mathcal{R}^d \mapsto \mathcal{R}^{d\times d}$, $m : \mathcal{R}^d \times \mathcal{R}^s \times A \mapsto \mathcal{R}^s$, $\sigma : \mathcal{R}^d \times \mathcal{R}^s \mapsto \mathcal{R}^{s\times s}$ are Lipschitz in the first and second (if any) arguments uniformly w.r.t. the third (if any);
- the least eigenvalues of $\gamma(z)\gamma(z)^T$, $\sigma(z, x)\sigma(z, x)^T$ are uniformly bounded away from zero (the *non-degeneracy* assumption);
- the initial values are fixed: $(z^\epsilon(0), x^\epsilon(0)) = (z_0, x_0)$;
- $B(\cdot), W(\cdot)$ are resp. d- and s-dimensional independent standard Brownian motions;
- $u(\cdot)$ is an A-valued control process with measurable paths satisfying the *non-anticipativity* condition: for $t \geq s$, $(B(t)-B(s), W(t)-W(s))$ is independent of

$$\hat{\mathcal{F}}_s := \text{ the completion of } \cap_{a\geq 0}\, \sigma(z^\epsilon(y), x^\epsilon(y), u(y), y \leq s + a).$$

We call such $u(\cdot)$ an admissible control. An important subclass of admissible controls is that of stationary Markov controls wherein $u(t)$ is of the form $\nu(z^\epsilon(t), x^\epsilon(t))$ for a measurable $\nu(z, x) : (z, x) \in \mathcal{R}^d \times \mathcal{R}^s \mapsto A$. Under such controls, $(z^\epsilon(\cdot), x^\epsilon(\cdot))$ will be a time-homogeneous Markov process. If it is stable, i.e., positive recurrent, it will have a unique stationary distribution thanks to our assumption of non-degeneracy (see, e.g., [1], Chapter 1).

The ergodic control problem is to minimize over all admissible $u(\cdot)$ the "ergodic cost"

$$\limsup_{t\uparrow\infty} \frac{1}{t} \int_0^t E[k(z^\epsilon(s), x^\epsilon(s), u(s))]ds. \tag{2.3}$$

Here the 'running cost' function $k : \mathcal{R}^d \times \mathcal{R}^s \times A \mapsto \mathcal{R}^+$ is continuous. We also assume the following:

(†) There exists an $\infty > M > 0$ such that for each $\epsilon \in (0, 1)$, the cost for at least one admissible control $u(\cdot)$ is $\leq M$.

As in [4], we shall work with the weak formulation of the above control problem[1] and assume that $u(\cdot)$ is a relaxed control. That is, for some compact metric space A', $A = \mathcal{P}(A')$:= the space of probability measures on A' with the Prohorov topology. Moreover, all functions above of the form $f(\cdots, u(t))$ (specifically, k and the components of h, m) are of the form $\int f'(\cdots, y)u(t, dy)$ (where $u(t) = u(t, dy)$) for an f' satisfying the same conditions as f except that the factor A of its domain is replaced by A'. This relaxation, originally introduced by L. C. Young in deterministic control, is a true relaxation in the sense that the attainable laws in the original set-up are dense in the set of attainable laws over relaxed controls (Corollary 2.3.6, pp. 53-54, [1]). As above, $\mathcal{P}(Z)$ for any Polish space Z will denote the Polish space of probability measures on Z with the Prohorov topology.

In view of our discussion in the introduction, we also introduce the *associated system*

$$dx(t) = m(z, x(t), u(t))dt + \sigma(z, x(t))dW(t), \; x(0) = x_0, \tag{2.4}$$

and the *averaged system*

$$dz(t) = \tilde{h}(z(t), \mu(t))dt + \gamma(z(t))dB(t), \; z(0) = z_0, \tag{2.5}$$

where

$$\tilde{h}(z, \mu) := \int h'(z, x, u)\mu(dx, du)$$

for $\mu \in \mathcal{P}(\mathcal{R}^s \times A')$. Then (2.4) is simply the dynamics of the fast component (2.2) with the slow component $z^\epsilon(t)$ frozen at a constant value z, followed by a time scaling $\frac{t}{\epsilon} \mapsto t$, whereas (2.5) is the dynamics of the slow component (2.1) with the dependence on control as well as the fast component averaged out with respect to a $\mathcal{P}(\mathcal{R}^s \times A')$-valued process $\mu(\cdot)$. Later on we shall see some natural candidates for $\mu(\cdot)$.

Define the *ergodic occupation measures* corresponding to the original, resp., associated and averaged systems, as follows.

(1) Define

$$\Phi^\epsilon_\nu(dz, dx, du) := \eta^\epsilon_\nu(dz, dx)\nu(du|z, x) \in \mathcal{P}(\mathcal{R}^d \times \mathcal{R}^s \times A'),$$

where η^ϵ_ν is the unique stationary distribution, if it exists, for (2.1)-(2.2) under the A-valued stationary Markov control ν. Let $\mathcal{G}^\epsilon$ denote the set of all such Φ^ϵ_ν.

(2) Define

$$\hat{\Phi}^z_{\nu'}(dx, du) := \zeta^z_{\nu'}(dx)\nu'(du|x) \in \mathcal{P}(\mathcal{R}^s \times A'),$$

where $\zeta^z_{\nu'}$ is the unique stationary distribution, if it exists, for (2.4) under the A-valued stationary Markov control ν'. Let $\mathcal{G}_z$ denote the set of all such $\hat{\Phi}^z_{\nu'}$.

[1]This allows us to assume without loss of generality that γ, σ are square matrices, because if they were not we could replace them by symmetric positive definite square-roots of $\gamma\gamma^T, \sigma\sigma^T$ resp. without any loss of generality (Section 5.3, [11]).

(3) Define

$$\tilde{\Phi}_{\tilde{\nu}}(dz, dx, du) := \beta_{\tilde{\nu}}(dz)\tilde{\nu}(dx, du|z) \in \mathcal{P}(\mathcal{R}^d \times \mathcal{R}^s \times A'),$$

where $\beta_{\tilde{\nu}}$ is the unique stationary distribution, if it exists, for (2.5) under stationary Markov control $\tilde{\nu} : z \in \mathcal{R}^s \mapsto \mathcal{G}_z$. Let $\tilde{\mathcal{G}}$ denote the set of all such $\tilde{\Phi}_{\tilde{\nu}}$.

For the averaged system, $\mathcal{G}_z$ will later serve as a state-dependent action space. When this interpretation is operative, in order to conform with our notation for action space, we shall use $\mathcal{A}_z$ for $\mathcal{G}_z$. The two will be thus used interchangeably depending on the context.

Clearly, the respective stationary Markov controls $\nu, \nu', \tilde{\nu}$ must be stable for the stationary distributions to exist. Furthermore, the cost (2.3) under stable stationary Markov control $\nu : \mathcal{R}^d \times \mathcal{R}^s \mapsto A$ is given by the integral

$$\int k' d\Phi^{\epsilon}_{\nu}.$$

We now summarize the results of [4] (see also [1], section 4.3). We introduce the following differential operators: Let

$$a(z, x) := \sigma\sigma^T(z, x) = [[a_{ij}(z, x)]]_{1 \le i,j \le s}.$$

Let ∇_y, ∇^2_y denote resp. the gradient and the Hessian w.r.t. the variable y. For $n \ge 1$, denote by $C(\mathcal{R}^n), C_b(\mathcal{R}^n), C_0(\mathcal{R}^n)$ resp. the space of continuous maps $\mathcal{R}^n \mapsto \mathcal{R}$, bounded continuous maps $\mathcal{R}^n \mapsto \mathcal{R}$, and continuous maps $\mathcal{R}^n \mapsto \mathcal{R}$ vanishing at infinity. Denote by $C^2(\mathcal{R}^n)$ the space of twice continuously differentiable maps $\mathcal{R}^n \mapsto \mathcal{R}$ and by $C^2_0(\mathcal{R}^n)$ the subset of $C^2(\mathcal{R}^n)$ of functions vanishing at ∞ along with its first and second order partial derivatives.

(1) Define $\mathcal{L}^u_z : C^2(\mathcal{R}^s) \mapsto C(\mathcal{R}^s \times A')$ by

$$\mathcal{L}^u_z(f)(x) := \frac{1}{2}\mathrm{tr}\left(a(z, x)\nabla^2_x f(x)\right) + \langle \nabla_x f(z, x), m'(z, x, u)\rangle, \ f \in C^2(\mathcal{R}^s).$$

(2) Define $\hat{\mathcal{L}}^u_\epsilon : C^2(\mathcal{R}^d \times \mathcal{R}^s) \mapsto C(\mathcal{R}^d \times \mathcal{R}^s \times A')$ by: for $f \in C^2(\mathcal{R}^d \times \mathcal{R}^s)$ with $f_z(\cdot) := f(z, \cdot) \in C^2(\mathcal{R}^s)\ \forall z$,

$$\begin{aligned}\hat{\mathcal{L}}^u_\epsilon(f)(z, x) \ &:= \ \frac{1}{2}\mathrm{tr}\left(\gamma(z)\gamma^T(z)\nabla^2_z f(z, x)\right) + \\ &\quad \langle \nabla_z f(z, x), h'(z, x, u)\rangle + \frac{1}{\epsilon}\mathcal{L}^u_z f_z(x).\end{aligned}$$

(3) Define $\tilde{\mathcal{L}}^\mu : C^2(\mathcal{R}^d) \mapsto C(\mathcal{R}^d), \mu \in \mathcal{A}_z$ by

$$\tilde{\mathcal{L}}^\mu f(z) := \frac{1}{2}\mathrm{tr}\left(\gamma(z)\gamma^T(z)\nabla^2_z f(z)\right) + \langle \nabla_z f(z), \tilde{h}(z, \mu)\rangle.$$

Observe that the three operators defined above, viz., $\hat{\mathcal{L}}^u_\epsilon, \mathcal{L}^u_z, \tilde{\mathcal{L}}^\mu$ are resp. the controlled extended generators for (2.1)-(2.2), (2.4) and (2.5). We shall define (relaxed) stationary Markov controls correspondingly, as measurable maps $x \in \mathcal{R}^s \mapsto A$ for the associated system and measurable maps $z \in \mathcal{R}^d \mapsto \mathcal{A}_z$ for the averaged system. They are stable if the corresponding controlled diffusion, perforce a time-homogeneous Markov process, is positive recurrent.

The following is immediate from Theorem 2.1 of [3].

Lemma 2.1. *(i) The set $\mathcal{G}^\epsilon$ is characterized by*

$$\begin{aligned} \mathcal{G}^\epsilon &= \{\Phi \in \mathcal{P}(\mathcal{R}^d \times \mathcal{R}^s \times A') : \\ &\qquad \int \hat{\mathcal{L}}^u_\epsilon f(z,x) d\Phi(dz,dx,du) = 0 \ \forall \ f \in C_0^2(\mathcal{R}^d \times \mathcal{R}^s)\}, \end{aligned} \tag{2.6}$$

(ii) The set $\mathcal{G}_z$ is characterized by

$$\mathcal{G}_z = \{\Phi \in \mathcal{P}(\mathcal{R}^s \times A') : \int \mathcal{L}^u_z f(x) d\Phi(dx,du) = 0 \ \forall \ f \in C_0^2(\mathcal{R}^s)\}, \tag{2.7}$$

(iii) The set $\tilde{\mathcal{G}}$ is characterized by

$$\begin{aligned} \tilde{\mathcal{G}} &= \{\Phi(dz,dx,du) := \beta_\mu(dz)\mu(dx,du|z) \in \mathcal{P}(\mathcal{R}^d \times \mathcal{R}^s \times A') : \\ &\qquad \mu \in \mathcal{G}_z, \ \int \tilde{\mathcal{L}}^{\mu(\cdot|z)} f(x) d\beta_\mu(x) = 0 \ \forall \ f \in C_0^2(\mathcal{R}^d)\}. \end{aligned} \tag{2.8}$$

Define

$$\tilde{k}(z,\mu) := \int k'(z,x,u)\mu(dx,du).$$

Here μ is as above, but is required to take values in the set $\mathcal{A}_z$, now viewed as a state-dependent action space for the controlled diffusion (2.5). The ergodic control problem for the averaged system, i.e., our candidate limiting problem, is then to minimize

$$\limsup_{t\uparrow\infty} \frac{1}{t} \int_0^t E\left[\tilde{k}(z(s),\mu(s))\right] ds \tag{2.9}$$

for $z(t) \in \mathcal{R}^d, \mu(t) \in \mathcal{A}_{z(t)}, t \geq 0$, as in (2.5). Under a stable stationary Markov control $\tilde{\nu} : z \in \mathcal{R}^d \mapsto \tilde{\nu}(z) \in \mathcal{A}_z$, (2.9) will equal $\int k' d\tilde{\Phi}_{\tilde{\nu}}$.

Our objective is to show that the averaged problem described by the controlled dynamics (2.5) with state-dependent action spaces $\mathcal{A}_z, z \in \mathcal{R}^d$, and cost (2.9) is a valid approximation to our original two time-scale ergodic control problem in the $\epsilon \downarrow 0$ limit.

Next we summarize the results of [4]. We assume the following throughout:

(††) There exist $V \in C^2(\mathcal{R}^s), g \in C(\mathcal{R}^d \times \mathcal{R}^s)$ such that

- $\lim_{\|x\|\uparrow\infty} V(x) = \infty$,
- $\lim_{\|x\|\uparrow\infty} g(z,x) = \infty$ uniformly in z in any compact set, and,

- for

$$\mathcal{L}f(z,x,u) \;:=\; \frac{1}{2}\mathrm{tr}\left(a(z,x)\nabla_x^2 f(x)\right) + \langle \nabla_x f(x), m'(z,x,u)\rangle, \qquad f \in C_0^2(\mathcal{R}^s),$$

V satisfies

$$\mathcal{L}V(z,x,u) < -g(z,x).$$

Say that k is *near-monotone* if

$$\liminf_{\|(z,x)\|\uparrow\infty} \inf_u k(z,x,u) = \infty. \tag{2.10}$$

In [4], three distinct cases are considered as described below. Here v^* is as in Theorem 2.2(ii) that follows.

(1) *The affine case with near-monotonicity:* Assume k to be near-monotone. In addition, assume:
  (a) $A' \subset \mathcal{R}^m$ for some $m \geq 1$ and is compact, with $h'(z,x,\cdot), m'(z,x,\cdot)$ affine and $k'(z,x,\cdot)$ strictly convex,
  (b) $\|h'(z,x,u)\| = o(k'(z,x,u))$ as $\|(z,x)\| \uparrow \infty$ and

$$\sup_u \|k'(z,x,u)\|^{1+a} \leq K(1+|g(z,x)|) \quad \forall z,x$$

  for some $K, a > 0$ and g as in (††).
  (c) There exists an $\epsilon_0 > 0$ such that for all $\epsilon \in (0,\epsilon_0)$, if μ^* is as in Theorem 2.2(ii) with

$$\mu^*(dx,du|z) = \zeta(dx|z)v^*(du|z,x),$$

  then as ϵ varies over $(0,\epsilon_0)$, v^* is a stable stationary Markov control for (2.1)-(2.2) and the corresponding stationary distributions are tight.

(Note that v^* is already stable for the averaged system. Here we require the stability to continue to hold for small $\epsilon > 0$.)

(2) *The general near-monotone case:* Assume k to be near-monotone. Define smooth approximations $v_\delta^*, \delta \in (0,1)$, to the v^* above by

$$\int f v_\delta^*(du|z,x) := \int\int f v^*(du|z-z', x-x')\psi_\delta(z',x')dz'dx' \tag{2.11}$$

for f in a countable convergence determining class in $C(A)$ and $\psi_\delta :=$ compactly supported smooth approximations of the Dirac measure, i.e.,

$$\psi_\delta(z'-z, x'-x)dx' \overset{\delta\downarrow 0}{\to} \delta_{(z,x)}(dz',dx') \text{ in } \mathcal{P}(\mathcal{R}^d \times \mathcal{R}^s).$$

We assume that for some $\delta_0, \epsilon_0, a > 0$, and for all $\delta \in (0,\delta_0]$ and $\epsilon \in (0,\epsilon_0]$, v_δ^* are stable stationary Markov controls for (2.1)-(2.2) and

$$\sup_u |k'(z^\epsilon(t), x^\epsilon(t), u)|^{1+a}$$

has a uniformly bounded stationary expectation under v^*_δ.

(3) *The stable case:* Assume that for some bounded open $B \subset \mathcal{R}^d \times \mathcal{R}^s$ and $\epsilon_0 > 0$ the following holds: for each $\epsilon \in (0, \epsilon_0)$, there exist $\Delta_\epsilon, a_\epsilon > 0$ and $V_1, V_2 \in C^2(\mathcal{R}^d \times \mathcal{R}^s)$ such that

$$\hat{\mathcal{L}}^u_\epsilon V^\epsilon_1 \leq -\Delta_\epsilon,\ \hat{\mathcal{L}}^u_\epsilon V^\epsilon_2 \leq -a_\epsilon V^\epsilon_1$$

for $(z, x) \notin B$.

A possible relaxation of (2.10) is pointed out in [4], equation (28).

Remark: While the use of v^*, whose existence is a part of the conclusions of Theorem 2.3, in the *assumptions* for the same theorem may appear self-referential, a look at [4] shows that the said existence follows purely from the preceding components of the above hypotheses, viz., near-monotonicity or the coupled Liapunov condition in the stable case.

Under any of the above sets of hypotheses, the following is established in [4] (see also [1], section 4.3):

Theorem 2.2. (i) *The ergodic control problem defined by (2.1), (2.2), (2.3) has an optimal stable A-valued stationary Markov control $v^*_\epsilon(du|z, x)$ with optimal cost $\int k' d\Phi^*_\epsilon$, where $\Phi^*_\epsilon \in \mathcal{G}^\epsilon$ is the corresponding optimal ergodic occupation measure.*

(ii) *The ergodic control problem defined by (2.5), (2.9) has an optimal stable stationary Markov control $\mu^*(dx, du|z) = \zeta^*(dx)v^*(du|x)$ with optimal cost $\int k' d\Phi^*_0$, where $\Phi^*_0 \in \tilde{\mathcal{G}}$ is the corresponding optimal ergodic occupation measure.*

(iii)

$$\liminf_{\epsilon \downarrow 0} \int k' d\Phi^*_\epsilon \geq \int k' d\Phi^*_0. \tag{2.12}$$

Theorem 2.2 in turn can be strengthened to the following:

Theorem 2.3. *Under any of the conditions above,*

$$\lim_{\epsilon \downarrow 0} \int k' d\Phi^*_\epsilon = \int k' d\Phi^*_0. \tag{2.13}$$

We refer the reader to [4] or Section 4.3 of [1] for the (lengthy) details of Theorems 2.2 and 2.3. Suffices to say that the former is a straightforward consequence of the upper semi-continuity of the set valued map

$$\epsilon \in [0, \epsilon_0] \mapsto Argmin_{\Phi \in \mathcal{G}^\epsilon} \left(\int k' d\Phi \right), \tag{2.14}$$

where we set $\mathcal{G}^\epsilon = \tilde{\mathcal{G}}$ for $\epsilon = 0$. For Theorem 2.3, however, one needs in addition to exhibit an optimal element of $\mathcal{G}^0$ as a limit point of a sequence $\mu_n \in \mathcal{G}^{\epsilon(n)}$ for $0 < \epsilon(n) \downarrow 0$. This program is carried out in *ibid.* under the three alternative sets of hypotheses mentioned above.

3. Main results

We now describe the Stackelberg control problem. We shall assume that A' is of the form $A' = A'_1 \times A'_2$ where A'_1, A'_2 are compact metric spaces, denoting action spaces for resp. player 1 or the 'leader' operating on a slow time scale, and player 2 or the 'follower' operating on the fast time scale. Spaces $A_i := \mathcal{P}(A'_i), i = 1, 2$, are defined correspondingly. The A-valued control process $u(\cdot) = u^\epsilon(\cdot)$ can then be written as $u^\epsilon(\cdot) = [u_1^\epsilon(\cdot), u_2^\epsilon(\cdot)]$ where $u_i^\epsilon(\cdot)$ is A_i-valued for $i = 1, 2$. We shall also assume that k is near-monotone. For $f = m, h$ or k, we shall assume that f is of the form

$$f(z, x, [u_1, u_2]) = f_1(z, x, u_1) + f_2(z, x, u_2) \tag{3.1}$$

for f_1, f_2 satisfying the same assumptions as the ones stipulated for f.

The main distinction from the preceding case is in the information structure we impose. We suppose that player 1 observes only $z^\epsilon(\cdot), u_1^\epsilon(\cdot)$, whereas player 2 can observe everything, i.e., $z^\epsilon(\cdot), u_1^\epsilon(\cdot), x^\epsilon(\cdot), u_2^\epsilon(\cdot)$. We view $u^\epsilon(\cdot)$ as a random variable taking values in the space $\mathcal{U}$ of measurable functions $\mu : [0, \infty) \mapsto A$ with the coarsest topology that renders continuous the maps $\mu \mapsto \int_s^t g(y) \int f d\mu(y) dy$ for all $t > s \geq 0$, $g \in L_2[s, t]$, $f \in C_b(A')$. This space is compact metrizable ([1], pp. 50-51.) We may write $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$ where $\mathcal{U}_i, i = 1, 2$, are path spaces for $u_1^\epsilon(\cdot), u_2^\epsilon(\cdot)$ resp., topologized analogously. Both are compact metrizable.

We further restrict $u_1^\epsilon(\cdot)$ as follows. Let the processes $(z^\epsilon(\cdot), x^\epsilon(\cdot), u^\epsilon(\cdot))$ in (2.1)-(2.2) be defined canonically on the probability space := the path space

$$\Omega := C([0, \infty); \mathcal{R}^d) \times C([0, \infty); \mathcal{R}^s) \times \mathcal{U},$$

equipped with the product Borel σ-field Ψ and probability measure $P :=$ the law of $(z^\epsilon(\cdot), x^\epsilon(\cdot), u^\epsilon(\cdot))$. Let $\Psi_t, t \geq 0$, denote the natural filtration of these processes, i.e., $\Psi_t :=$ the P-completion of $\cap_{s>t}\sigma(z^\epsilon(y), x^\epsilon(y), u^\epsilon(y), y \leq s)$ for each $t \geq 0$. Define a new probability measure Q on (Ω, Ψ) as follows: $Q_t := Q|_{\Psi_t}$ is mutually absolutely continuous w.r.t. $P_t := P|_{\Psi_t}$ for all $t \geq 0$ with

$$\begin{aligned} \Lambda_t &:= \frac{dP_t}{dQ_t} \\ &= exp\Big(\int_0^t \langle \gamma(z^\epsilon(s))^{-1} h(z^\epsilon(s), x^\epsilon(s), u^\epsilon(s))), d\bar{B}(s) \rangle \\ &\quad - \frac{1}{2} \int_0^t \| \gamma(z^\epsilon(s))^{-1} h(z^\epsilon(s), x^\epsilon(s), u^\epsilon(s)) \|^2 ds \Big), \end{aligned} \tag{3.2}$$

where

$$\bar{B}(t) := B(t) + \int_0^t \gamma(z^\epsilon(s))^{-1} h(z^\epsilon(s), x^\epsilon(s), u^\epsilon(s)) ds$$

is a d-dimensional standard Brownian motion independent of $W(\cdot)$ under Q by the Girsanov theorem ([11], section 6.4). Then under Q, $z^\epsilon(\cdot)$ satisfies

$$dz^\epsilon(t) = \gamma(z^\epsilon(t))d\bar{B}(t).$$

We shall say that $u_1^\epsilon(\cdot)$ is *wide sense admissible* [6] if under Q, $z^\epsilon(t) - z^\epsilon(s)$ is independent of $\{u_1^\epsilon(y), z^\epsilon(y), \bar{B}(y), y \leq s; W(\cdot), x^\epsilon(0)\}$ for all $t > s \geq 0$. This includes in particular the controls adapted to the natural filtration of $z^\epsilon(\cdot)$, i.e., the Q-completion of $\cap_{s>t}\sigma(z^\epsilon(y), y \leq s)$, $t \geq 0$. These are the so called strict sense admissible controls. Wide sense admissible controls is a relaxation of this notion and has the advantage that the laws of such $(\bar{B}(\cdot), u_1^\epsilon(\cdot))$ under Q form a convex compact set in $(\mathcal{P}(C([0,\infty); \mathcal{R}^d) \times \mathcal{U})$ (*ibid.*). Furthermore, the property of wide sense admissibility is preserved under convergence in law (*ibid.*), defined as it is in terms of *independence* rather than *conditional independence*. Finally, the laws under strict sense admissible controls are dense in the set of laws under wide sense admissible controls, making it a legitimate relaxation. See [6] for details.

We assume throughout what follows that $(\dagger), (\dagger\dagger)$ continue to hold. Let $\mathcal{F}_t^\epsilon :=$ the completion of $\cap_{t'>t}\sigma(z^\epsilon(s), u_1^\epsilon(s), s \leq t')$ for $t \geq 0$. Define the $\mathcal{P}(\mathcal{R}^s \times A_2')$-valued process λ_t^ϵ for $t \geq 0$ by:

$$\int f(x,y)\lambda_t^\epsilon(dx,dy) := E\left[\int f(x^\epsilon(t),y)u_2^\epsilon(t)(dy)\Big|\mathcal{F}_t^\epsilon\right],$$

for $f \in$ a suitable countable convergence determining class in $C_b(\mathcal{R}^s \times A_2')$. Let $\pi_t^\epsilon(dx) := \lambda_t^\epsilon(dx, A_2')$, which then is the regular conditional law of $x^\epsilon(t)$ given $\mathcal{F}_t^\epsilon$. Our approach is based on the following lemmas. In what follows, we introduce the notation

$$\breve{h}(z,x,\mu,w) := \int h'(z,x,[w',w])\mu(dw'),\ (z,x,\mu,w) \in \mathcal{R}^d \times \mathcal{R}^s \times A_1 \times A_2'.$$

Correspondingly, we also define

$$\mathcal{L}_z^{[u,w]} := \frac{1}{2}\mathrm{tr}\left(a(z,x)\nabla_x^2 f(x)\right) + \langle \nabla_x f(z,x), \int m'(z,x,[w',w])u(dw')\rangle.$$

Lemma 3.1. *Equation (2.1) can be rewritten as*

$$dz^\epsilon(t) = \int \breve{h}(z^\epsilon(t),x,u_1^\epsilon(t),y)\lambda_t^\epsilon(dx,dy)dt + \gamma(z^\epsilon(t))d\tilde{B}(t), \tag{3.3}$$

where for $t \geq 0$,

$$\tilde{B}(t) = B(t) + \int_0^t \Big(\gamma(z^\epsilon(s))^{-1}\Big(h(z^\epsilon(s),x^\epsilon(s),u^\epsilon(s)) - \int \breve{h}(z^\epsilon(s),x,u_1^\epsilon(s),y)\lambda_s^\epsilon(dx,dy)\Big)\Big)ds$$

is a d-dimensional standard Brownian motion independent of $W(\cdot)$.

The proof is immediate from Levy's martingale characterization of Brownian motion (see, e.g., [14], Theorem 4.2). Write $\pi_t^\epsilon(f) := \int f d\pi_t^\epsilon$ for any bounded measurable $f : \mathcal{R}^s \mapsto \mathcal{R}$. Similarly, $\lambda_t^\epsilon(g) := \int g d\lambda_t^\epsilon$ for bounded measurable $g : \mathcal{R}^s \times A_2' \mapsto \mathcal{R}$.

Lemma 3.2. *The evolution of* $\{\pi_t^\epsilon\}$ *as a probability measure-valued process is given by the equations of nonlinear filtering:*

$$\begin{aligned} d\pi_t^\epsilon(f) \quad = \quad & \frac{1}{\epsilon}\lambda_t^\epsilon(\mathcal{L}_{z^\epsilon(t)}^{[u_1^\epsilon(t),\cdot]} f)dt + \\ & \Big\langle \lambda_t^\epsilon(f(\cdot)\gamma^{-1}(z^\epsilon(t))\breve{h}(z^\epsilon(t),\cdot,u_1^\epsilon(t),\cdot)) - \\ & \pi_t^\epsilon(f)\lambda_t^\epsilon\Big(\gamma^{-1}(z^\epsilon(t))\breve{h}(z^\epsilon(t),\cdot,u_1^\epsilon(t),\cdot)\Big), d\tilde{B}(t)\Big\rangle, \end{aligned} \tag{3.4}$$

for $t \geq 0$.

This follows along standard lines, see, e.g., [1], section 8.1. We sketch the details in the Appendix. The one difference with the usual set up is the appearance of the process $\lambda_\cdot^\epsilon$ on the right because of an unobserved (by the leader) control process. This causes only a minor change in the derivation. Note that we are not interested in the uniqueness of the solution to (3.4) as it will not be required for our purposes.

For $\epsilon \in (0,1)$, let $(\breve{z}^\epsilon(\cdot), \breve{x}^\epsilon(\cdot), \breve{u}^\epsilon(\cdot))$ be a jointly stationary solution to (2.1)-(2.2) that is optimal. From now on we view these processes as being defined over the entire time axis $\mathcal{R}$. Let $\breve{\pi}_t^\epsilon, \breve{\lambda}_t^\epsilon$ denote the corresponding conditional laws given $\mathcal{F}_t^\epsilon,\ t \geq 0$, featuring in the nonlinear filter. These will also be stationary. We view $\breve{z}^\epsilon(\cdot), \breve{x}^\epsilon(\cdot)$ as resp. $C(\mathcal{R};\mathcal{R}^d), C(\mathcal{R};\mathcal{R}^s)$ valued random variables, where these spaces are endowed with the topology of uniform convergence on compact intervals. We view $\breve{u}_1^\epsilon(\cdot)$ as a random element of $\widetilde{\mathcal{U}}_1 :=$ the space of measurable paths $\mathcal{R} \mapsto A_1$ with the coarsest topology that renders continuous the maps $\kappa(\cdot) \in \mathcal{U} \mapsto \int_s^t g(y) \int f d\kappa(y) dy$ for all $t > s$ in $\mathcal{R}$, all $g \in L_2[s,t]$, and all $f \in C(A_1')$. Likewise, let $\widetilde{\mathcal{R}}^s = \mathcal{R}^s \cup \{\infty\}$ denote the one point compactification of $\mathcal{R}^s$ and view $\breve{\lambda}_\cdot^\epsilon$ as a random element of $\mathcal{V} :=$ the space of measurable paths $\mathcal{R} \mapsto \mathcal{P}(\widetilde{\mathcal{R}}^s \times A_2')$ with the coarsest topology that renders continuous the maps $\zeta(\cdot) \in \mathcal{V} \mapsto \int_s^t g(y) \int f d\zeta(y) dy$ for all $t > s$ in $\mathcal{R}$, all $g \in L_2[s,t]$, and all $f \in C(\widetilde{\mathcal{R}}^s \times A_2')$. Define $\mathcal{V}_0$ likewise with $\mathcal{R}^s$ replacing $\widetilde{\mathcal{R}}^s$. Both $\widetilde{\mathcal{U}}_1$ and $\mathcal{V}$ are compact metrizable and hence Polish (see, e.g., Theorem 2.3.1, pp. 50-51 of [1]).

Remark: In the above definition of $\mathcal{V}$, it suffices to consider $f \in C(\tilde{\mathcal{R}}^s \times A_2')$ such that $f|_{\mathcal{R}^s \times A_2'} \in$ a countable convergence determining class in

$$C_0(\mathcal{R}^s \times A_2') := \{f' \in C(\mathcal{R}^s \times A_2') : \lim_{\|x\|\uparrow\infty} \max_{u \in A_2'} |f(x,u)| = 0\}$$

(e.g., a countable dense set in the unit ball of $C_0(\mathcal{R}^s \times A_2')$ around the origin), and $f(\infty, \cdot) \equiv 0$.

Lemma 3.3. *The laws of* $(\breve{z}^\epsilon(\cdot), \breve{\lambda}^\epsilon_\cdot, \breve{u}^\epsilon(\cdot))$ *are tight in* $\mathcal{P}(C(\mathcal{R};\mathcal{R}^d) \times \mathcal{V} \times \mathcal{U})$.

Proof. Since $\mathcal{U}, \mathcal{V}$ are compact, we need only verify that the laws of $\breve{z}^\epsilon(\cdot)$ are tight. By (†),

$$E[k(\breve{z}^\epsilon(t), \breve{x}^\epsilon(t), \breve{u}^\epsilon(t))] \leq M < \infty.$$

In view of our near-monotonicity assumption regarding k, the marginal laws of $\breve{z}^\epsilon(t)$, which do not depend on t, remain tight over $\epsilon \in (0,1)$.) For $f \in C_0^2(\mathcal{R}^d)$,

$$f(\breve{z}^\epsilon(t)) - \int_0^t \hat{\mathcal{L}}_\epsilon^{\breve{u}^\epsilon(s)} f(\breve{z}^\epsilon(s), \breve{x}^\epsilon(s))ds, \quad t \geq 0, \tag{3.5}$$

is a martingale w.r.t. $\{\mathcal{F}_t^\epsilon\}$. Applying Theorem 9.4, p. 145, [5], we get tightness of the laws of $\breve{z}^\epsilon(\cdot)$. □

Let $(\breve{z}^*(\cdot), \breve{\lambda}^*_\cdot, \breve{u}^*(\cdot))$ denote a subsequential limit in law of $(\breve{z}^\epsilon(\cdot), \breve{\lambda}^\epsilon_\cdot, \breve{u}^\epsilon(\cdot))$ along some $\epsilon = \epsilon(n) \downarrow 0$. Let $\breve{\pi}_t^*$ denote the marginal of $\breve{\lambda}_t^*$ on $\mathcal{R}^s$ for $t \in \mathcal{R}$. These limit processes will be jointly stationary, being limits in law of jointly stationary processes.

Lemma 3.4. $\breve{\pi}_t^*(\mathcal{R}^s) = 1 \ \forall t$ *a.s.*

Proof. Note that for $\epsilon > 0$,

$$\breve{\pi}_t^\epsilon(\mathcal{R}^s) = 1 \text{ a.s. } \forall\ t.$$

On the other hand,

$$E\left[\int f d\breve{\pi}_t^\epsilon\right] = E\left[f(x^\epsilon(t))\right] \ \forall t.$$

We claim that the laws μ_t^ϵ of $x^\epsilon(t), 0 < \epsilon < \epsilon_0,\ t \geq 0$, are tight. To see this, first note that the marginals do not change with time scaling, so we need only prove tightness of the marginals of the time-scaled equation

$$d\tilde{x}^\epsilon(t) = m(\tilde{z}^\epsilon(t), \tilde{x}^\epsilon(t), \tilde{u}(t))dt + \sigma(\tilde{z}^\epsilon(t), \tilde{x}^\epsilon(t))d\tilde{W}(t),$$

where tilde denotes the time-scaled version after the time scaling $\frac{t}{\epsilon} \mapsto t$. Treating $\tilde{z}^\epsilon(\cdot)$ as a 'control process' taking values in $\mathcal{R}^d$ and using (††) in conjunction with the implication $(iv) \implies (i) \implies (iii) \implies (vi)$[2] of Lemma 3.3.4, pp. 97-98, [1], the tightness claim follows. But $\int f d\mu_t^\epsilon = E\left[\int f d\pi^\epsilon(t)\right] \ \forall\ f : \mathcal{R}^s \mapsto \mathcal{R}$ bounded measurable. Hence by Lemma 8.3.1, pp. 286-287, [1], $\breve{\pi}_t^\epsilon, 0 < \epsilon < \epsilon_0, t \geq 0$, are tight as $\mathcal{P}(\mathcal{R}^s)$-valued random variables, leading to $\breve{\pi}_t^*(\mathcal{R}^s) = 1$. By stationarity, this claim extends to all $t \in \mathcal{R}$. □

Corollary 3.5. *The laws of* $(\breve{z}^\epsilon(\cdot), \breve{\lambda}^\epsilon_\cdot, \breve{u}^\epsilon(\cdot))$ *are tight in* $\mathcal{P}(C(\mathcal{R};\mathcal{R}^d) \times \mathcal{V}_0 \times \mathcal{U})$.

Proof. This is immediate from the above lemma and the remark preceding Lemma 3.3. □

[2]This chain of implications can be worked out without the compactness of control space.

We shall need the following notation. For a random process $q(\cdot)$, let $q([s,s'])$ for $s < s'$ in $\mathcal{R}$ denote the trajectory of the process $q(\cdot)$ restricted to $[s,s']$, viewed as a random element of the corresponding path space (e.g., $C(\mathcal{R},\mathcal{R}^d)$ or $\mathcal{U}$) restricted to $[s,s']$ with the corresponding topology.

Lemma 3.6. *The processes* $(\breve{z}^*(\cdot), \breve{\lambda}^*_\cdot, \breve{u}^*(\cdot))$ *satisfy*

$$d\breve{z}^*(t) = \breve{\lambda}^*_t(\breve{h}(\breve{z}^*(t),\cdot,[\breve{u}^*_1(t),\cdot]))dt + \gamma(\breve{z}^*(t))d\breve{B}(t), \tag{3.6}$$

for a d-dimensional standard Brownian motion $\breve{B}$.

Proof. For any $f \in C^2_0(\mathcal{R}^d)$, (3.5) implies that

$$f(\breve{z}^\epsilon(t)) - \int_0^t \breve{\lambda}^\epsilon_s \left(\tilde{\mathcal{L}}^{[\breve{u}^\epsilon_1(s),\cdot]}_{\breve{z}^\epsilon(s)} f(\cdot)\right) ds, \quad t \geq 0,$$

is a martingale w.r.t. $\{\mathcal{F}^\epsilon_t\}$. Equivalently, fot $t > t_0$,

$$E\left[\left(f(\breve{z}^\epsilon(t)) - f(\breve{z}^\epsilon(t_0)) - \int_{t_0}^t \breve{\lambda}^\epsilon_s \left(\tilde{\mathcal{L}}^{[\breve{u}^\epsilon_1(s),\cdot]}_{\breve{z}^\epsilon(s)} f(\cdot)\right) ds\right) g(z^\epsilon([0,t_0]),\breve{u}^\epsilon_1([0,t_0]))\right] = 0, \tag{3.7}$$

for any bounded continous g on $C([0,t_0];\mathcal{R}^d) \times \mathcal{U}_{1t_0}$, where $\mathcal{U}_{1t_0}$ is defined and topologized exactly the same way as $\mathcal{U}_1$ except that the underlying time interval is restricted to $[0,t_0]$. The relation (3.7) is preserved under convergence in law. Hence by passing to an appropriate subsequential limit as $\epsilon \downarrow 0$, we have that

$$f(\breve{z}^*(t)) - \int_0^t \breve{\lambda}^*_s \left(\tilde{\mathcal{L}}^{[\breve{u}^*_1(s),\cdot]}_{\breve{z}^*(s)} f\right) ds, \quad t \geq 0,$$

is a martingale w.r.t. $\{\mathcal{F}^0_t\}$, where

$$\mathcal{F}^0_t := \text{ the completion of } \cap_{t'>t}\, \sigma(\breve{z}^*(s),\breve{u}^*_1(s), s \leq t').$$

The claim now follows by standard martingale representation theorems, see, e.g., [14]. □

Lemma 3.7. *The control process* $u^*_1(\cdot)$ *is wide sense admissible.*

Proof. (Sketch) This is proved by a standard argument, essentially adapted from [6]. Wide sense admissibility of $\breve{u}^\epsilon_1(\cdot)$ is equivalent to the statement that, for $t_2 > t_1 > t_0$ in $\mathcal{R}$ and $f \in C_b(\mathcal{R}^d)$,

$$\begin{aligned} & E_0\left[f(\breve{z}^\epsilon(t_2) - \breve{z}^\epsilon(t_1))g(\breve{u}^\epsilon([t_0,t_1]),\breve{z}^\epsilon([t_0,t_1]),\bar{B}([t_0,t_1]),W(\cdot))\right] \\ = \; & E_0\left[f(\breve{z}^\epsilon(t_2) - \breve{z}^\epsilon(t_1))\right] E\left[g(\breve{u}^\epsilon([t_0,t_1]),\breve{z}^\epsilon([t_0,t_1]),\bar{B}([t_0,t_1]),W(\cdot))\right] \end{aligned}$$

for any bounded continuous g on the relevant product space. We need to prove that this continues to hold in the $\epsilon \downarrow 0$ limit. This can be done as follows. First rewrite the equation as

$$\begin{aligned} & E\left[f(\check{z}^\epsilon(t_2) - \check{z}^\epsilon(t_1))g(u^\epsilon([t_0,t_1]), \check{z}^\epsilon([t_0,t_1]), \bar{B}([t_0,t_1]), W(\cdot))\widetilde{\Lambda}^\epsilon\right] \\ = \; & E\left[f(\check{z}^\epsilon(t_2) - \check{z}^\epsilon(t_1))\widetilde{\Lambda}^\epsilon\right] E\left[g(u^\epsilon([t_0,t_1]), \check{z}^\epsilon([t_0,t_1]), \bar{B}([t_0,t_1]), W(\cdot))\widetilde{\Lambda}^\epsilon\right], \end{aligned}$$

where

$$\begin{aligned} \widetilde{\Lambda}^\epsilon \; := \; & exp\Big(- \int_{t_0}^{t_2} \langle \gamma(\check{z}^\epsilon(s))^{-1} \int \check{h}(\check{z}^\epsilon(s), x, \check{u}_1^\epsilon(s), y)\lambda_s^\epsilon(dx,dy), dB(s)\rangle \\ & - \frac{1}{2}\int_0^t \|\gamma(\check{z}^\epsilon(s))^{-1} \int \check{h}(\check{z}^\epsilon(s), x, \check{u}_1^\epsilon(s), y)\lambda_s^\epsilon(dx,dy)\|^2 ds\Big). \end{aligned} \tag{3.8}$$

Then to pass to the limit in law in order to claim the same for $\epsilon = 0$, we need to verify uniform integrability of $\widetilde{\Lambda}^\epsilon, \epsilon > 0$. This follows by a standard criterion of Portenko ([10], Chapter I). □

Next we characterize the process $\check{\lambda}^*_\cdot$. Let $\bar{\mathcal{G}}_{z,w}$ denote the set of ergodic occupation measures for the associated system redefined as

$$dx(t) = m(z, x(t), [w, u_2(t)])dt + \sigma(z, x(t))dW(t), \; x(0) = x_0. \tag{3.9}$$

That is, $\bar{\mathcal{G}}_{z,w}$ is the set of probability measures

$$\eta_{z,w}(dx, du) = \eta^1_{z,w}(dx)\eta^2_{z,w}(du|x)$$

such that under $u_2(\cdot) \equiv \eta^2_{z,w}(du|\cdot)$, (3.9) is positive recurrent with unique stationary distribution $\eta^1_{z,w}$.

Lemma 3.8. *The set valued map $(z,w) \mapsto \bar{\mathcal{G}}_{z,w}$ is a nonempty compact convex valued upper semicontinuous map and $\check{\lambda}^*_t \in \bar{\mathcal{G}}_{\check{z}^*(t), \check{u}^*_1(t)}$ a.e.*

Proof. By Theorem 2.1 of [3], $\eta_{z,w}$ is uniquely characterized by:

$$\eta \in \bar{\mathcal{G}}_{z,w} \Longleftrightarrow \int \mathcal{L}_z^{[w,w']} f \eta_{z,w}(dx, dw') = 0 \;\; \forall \; f \in C_0^2(\mathcal{R}^s). \tag{3.10}$$

By (††), the $\eta_{z,w}$ remain tight as (z,w) vary over a compact subset. The first claim now follows from the fact that (3.10) is preserved under convex combinations and convergence in $\mathcal{P}(C([0,t_0];\mathcal{R}^s) \times \mathcal{U}_{2t_0})$, where $\mathcal{U}_{2t_0}$ is defined in a manner analogous to $\mathcal{U}_{1t_0}$. From (3.4), we have for $s \in \mathcal{R}, \delta > 0$,

$$\begin{aligned} & \epsilon(\check{\pi}^\epsilon_{s+\delta}(f) - \check{\pi}^\epsilon_s(f)) \\ & \int_s^{s+\delta} \check{\lambda}^\epsilon_t(\mathcal{L}^{[\check{u}^\epsilon_1(t),\cdot]}_{\check{z}^\epsilon(t)} f(\cdot))dt + \\ & \sqrt{\epsilon}\int_s^{s+\delta} \langle \check{\lambda}^\epsilon_t(f(\cdot)\gamma^{-1}(\check{z}^\epsilon(t))\check{h}(\check{z}^\epsilon(t), \cdot, \check{u}^\epsilon(t), \cdot)) - \end{aligned} \tag{3.11}$$

$$\breve{\pi}_t^\epsilon(f)\breve{\lambda}_t^\epsilon(\gamma^{-1}(\breve{z}^\epsilon(t))\breve{h}(\breve{z}^\epsilon(t),\cdot,\breve{u}^\epsilon(t),\cdot)), d\tilde{B}(t)\rangle. \tag{3.12}$$

Passing to the limit along $\epsilon = \epsilon(n) \downarrow 0$, we get

$$\int_s^{s+\delta} \lambda_t^\epsilon \left(\mathcal{L}_{\breve{z}^*(t)}^{[\breve{u}_1^*(t),\cdot]} f\right) dt = 0, \text{ a.s.}$$

Dividing by δ and letting $\delta \downarrow 0$, we get

$$\lambda_t^* \left(\mathcal{L}_{\breve{z}^*(t)}^{[\breve{u}_1^*(t),\cdot]} f\right) = 0 \text{ a.s.}$$

at all Lebesgue points, hence a.e. Considering f in a countable dense subset of $C_0^2(\mathcal{R}^s)$, this holds for all such f outside a common Lebesgue-null set, hence by a density argument, for all $f \in C_0^2(\mathcal{R}^s)$ as well. The claim now follows by (3.10). □

In summary, (3.6) can be replaced by

$$d\breve{z}^*(t) = \int \breve{h}(\breve{z}^*(t), x, \breve{u}_1^*(t), y)\eta_{\breve{z}^*(t),\breve{u}_1^*(t)}(dx, dy)dt + \gamma(\breve{z}^*(t))d\breve{B}(t). \tag{3.13}$$

We now redefine the averaged dynamics as

$$dz(t) = \check{h}(z(t), [u(t), \tilde{u}(t)])dt + \gamma(z(t))dB(t) \tag{3.14}$$

with $\tilde{u}(t) \in \bar{\mathcal{G}}_{z(t),u(t)}$. The notation here is as follows. We treat $[u(\cdot), \tilde{u}(\cdot)]$ as a measurable non-anticipative control process taking values in the state-dependent action space $\mathcal{A}_z^* := \cup_{w \in A_1'}(\{w\} \times \bar{\mathcal{G}}_{z,w})$ with relative topology inherited from $A_1' \times \mathcal{P}(\mathcal{R}^d \times A_2')$. By Lemma 3.7, this is a compact space for each z and the map $z \mapsto \mathcal{A}_z^*$ is upper semicontinuous. In turn, the map $\check{h} : \cup_{z \in \mathcal{R}^d}(\{z\} \times \mathcal{A}_z^*) \mapsto \mathcal{R}^d$ is defined by

$$\check{h}(z, [u, \tilde{u}]) = \int h'(z, x, [y_1, y_2])\tilde{u}(dx, dy_2)u(dy_1),$$

for $z \in \mathcal{R}^d, [u, \tilde{u}] \in \mathcal{A}_z^*$.

We consider this averaged system with the objective of minimizing

$$\limsup_{t \uparrow \infty} \frac{1}{t} \int_0^t E\left[\check{c}(z(s), [u(s), \tilde{u}(s)])\right] ds \tag{3.15}$$

where

$$\check{c}(z, [u, \tilde{u}]) = \int c'(z, x, [y_1, y_2])\tilde{u}(dx, dy_2)u(dy_1).$$

Let $v^* = [\hat{v}^*, \tilde{v}^*]$ denote an optimal stable relaxed Markov control for this problem. The existence of v^* follows as in [1], section 3.4, with minor modifications to account for the state dependence of the action space $\mathcal{A}_\cdot^*$.

We now state the counterpart of the '*general near-monotone case*' described in the preceding section for the single agent control problem. Define $v_\delta^* = [\hat{v}_\delta^*, \tilde{v}_\delta^*]$ for $\delta > 0$ as before, identifying $v^* = [\hat{v}^*, \tilde{v}^*]$ with $\delta = 0$. Write $\tilde{v}_\delta^*(dx, dy)$ as $\tilde{v}_\delta^*(dx, dy|z, w)$ for $z \in \mathcal{R}^d, w \in A_1'$ so as to render explicit its dependence on (z, w), given that it was chosen from the (z, w)-dependent set $\bar{\mathcal{G}}_{z,w}$. We take this dependence to be measurable using a standard measurable selection theorem, see, e.g., [13]. Disintegrate it further as

$$\tilde{v}_\delta^*(dx, dy|z, w) = \breve{v}_\delta(dy|z, x, w)\phi_\delta(dx|z, w)$$

for $\breve{v}_\delta : \mathcal{R}^d \times \mathcal{R}^s \times A_1' \mapsto A_2$ and $\phi_\delta : \mathcal{R}^d \times A_1' \mapsto \mathcal{P}(\mathcal{R}^s)$. Consider $\bar{v}_\delta^* = [\hat{v}^*, \breve{v}_\delta]$. Our assumption is as follows.

(† † †) For some $\delta_0, \epsilon_0, a > 0$, and for all $\delta \in (0, \delta_0]$, $\epsilon \in (0, \epsilon_0]$,

- $\bar{v}_\delta^*$ are stable stationary Markov controls for (2.1)-(2.2), and,
- $\sup_u |k(z^\epsilon(t), x^\epsilon(t), u)|^{1+a}$ has a uniformly bounded stationary expectation under $\bar{v}_\delta^*$.

Observe that the stationary Markov controls $\bar{v}_\delta^* = [\hat{v}^*, \breve{v}_\delta]$ have a very specific structure: $\hat{v}^*$ is allowed to depend only on the slow component of the state whereas $\breve{v}_\delta$ has no such restriction. This is in tune with the Stackelberg team problem we have formulated. Also note that the uniform boundedness of stationary expectations of $\sup_u |k(z^\epsilon(t), x^\epsilon(t), u)|^{1+a}$ under $\bar{v}_\delta^*$ implies in particular that the corresponding stationary distributions are tight.

Theorem 3.9. *In the general near-monotone case for the Stackelberg team problem, an optimal control for the averaged system exists and the corresponding optimal cost is the limit as $\epsilon \downarrow 0$ of the optimal costs of the ϵ-perturbed problems.*

Proof. Let $z^*(\cdot)$ denote the stationary optimal solution to (3.14) under v^* as above. Since $\bar{v}_\delta^*$ are stable for $\epsilon \in (0, \epsilon_0)$, $\delta \in (0, \delta_0)$, the corresponding stable stationary solutions $(z^{\epsilon,\delta}(\cdot), x^{\epsilon,\delta}(\cdot))$ exist and their stationary marginals are tight as observed above. An application of Theorem 9.4, p. 145, [5], as in the proof of Lemma 3.3 then gives us tightness of laws of $z^{\epsilon,\delta}(\cdot)$. Write the dynamics of $z^{\epsilon,\delta}(\cdot)$ as

$$dz^{\epsilon,\delta}(t) = \int h(z^{\epsilon,\delta}(t), x, [\hat{v}_\delta^*(z^{\epsilon,\delta}(t)), \breve{v}_\delta^*(x)])\pi_t^\epsilon(dx)dt + \gamma(z^{\epsilon,\delta}(t))d\tilde{B}(t),$$

where $\pi_t^{\epsilon,\delta}$ is the regular conditional law of $x^{\epsilon,\delta}(t)$ given $z^{\epsilon,\delta}(s), s \le t$. Letting $\epsilon \downarrow 0$, argue as in the proof of Lemma 3.6 to conclude that any limit point in law must be a stationary process satisfying

$$dz^{0,\delta}(t) = \int h(z^{0,\delta}(t), x, [\hat{v}_\delta^*(z^{0,\delta}(t)), \breve{v}_\delta^*(x)])\pi_t^{0,\delta}(dx)dt + \gamma(z^{0,\delta}(t))d\tilde{B}(t),$$

where we use the continuity of $\breve{v}_\delta^*(\cdot)$, with $\pi_\cdot^{0,\delta}$ being an appropriate subsequential limit in law of $\pi_\cdot^{\epsilon,\delta}$. Now argue as in Lemma 3.7, using continuity of $\bar{v}_\delta^*(\cdot)$, to conclude that $\pi_t^{0,\delta} = \tilde{\eta}_{z^{0,\delta}(t)}^*$ where $\tilde{\eta}_z^*(dx)$ is the unique stationary distribution of the process $x^{0,\delta}(\cdot)$ given by

$$dx^{0,\delta}(t) = m(x^{0,\delta}(t), z, [\hat{v}_\delta^*(z), \breve{v}_\delta^*(x^{0,\delta}(t))])dt + \sigma(z, x^{0,\delta}(t))dW(t). \tag{3.16}$$

In particular, $z^{0,\delta}(\cdot)$ is the stationary solution of the averaged system under $\bar{v}_\delta^*$. Since $z^{\epsilon,\delta}(\cdot), \epsilon \in (0, \epsilon_0), \delta \in (0, \delta_0)$ are tight, so are $z^{0,\delta}(\cdot), \delta \in (0, \delta_0)$. Now note that

$\bar{v}^*_\delta \to v^*$ in the 'topology of Markov controls' defined in [1], section 2.4. Thus we can argue as in *ibid.* to conclude that

$$(z^{\epsilon,\delta}(\cdot), v^*_\delta(z^\epsilon(\cdot))) \to (z^*(\cdot), v^*(z^*(\cdot)))$$

in law, i.e., in $\mathcal{P}(C(\mathcal{R}, \mathcal{R}^d) \times \mathcal{U})$.

Now let $\beta(\epsilon)$ denote the optimal cost for (2.1)-(2.2) in the Stackelberg case, with $\epsilon = 0$ being identified with the averaged system. Then Lemma 3.6 yields lower semicontinuity of the map $\epsilon \mapsto \beta(\epsilon)$ at $\epsilon = 0$ and the foregoing proves its upper semicontinuity. The claim follows. □

The claims for the affine case and the stable case are proved analogously. For the affine case, v^* is unique and continuous, so the argument is much simpler as in [4] or [1], section 4.3. In particular, the smoothing of v^* by ψ_δ is not needed. The stable case is also simpler because the issue of verifying tightness disappears altogether, simplifying the argument.

Remarks: 1. The condition (3.1) is necessitated by the topology for $\mathcal{U}_1 \times \mathcal{U}_2$ we work with, which gives continuity of $u_i(\cdot) \in \mathcal{U}_i \mapsto \int_0^t g(y) \int f du_i(y) dy \in \mathcal{R}$ for $t > s, g \in L_2[s,t], f \in C(A'_i)$ for $i = 1, 2$, separately, but not of $(u_1(\cdot), u_2(\cdot)) \in \mathcal{U} \mapsto \int_s^t g(y) \int f du_1(y) du_2(y) dy$ for t, s, g as above and $f \in C(A')$.

2. We have considered a team problem with a common cost function k. A genuine game formulation wherein the leader and the follower have differing cost criteria remains to be analyzed.

APPENDIX

We derive here the nonlinear filtering equation (3.4). Consider the processes $(z^\epsilon(\cdot), x^\epsilon(\cdot), u^\epsilon(\cdot) = [u^\epsilon_1(\cdot), u^\epsilon_2(\cdot)])$ in (2.1)-(2.2) to be defined canonically on the probability space (Ω, Ψ, P) with $\Omega :=$ the path space

$$C([0,\infty); \mathcal{R}^d) \times C([0,\infty); \mathcal{R}^s) \times \mathcal{U},$$

equipped with the product Borel σ-field Ψ and probability measure $P :=$ the law of $(z^\epsilon(\cdot), x^\epsilon(\cdot), u^\epsilon(\cdot))$. Let $\Psi_t, t \geq 0$, denote the natural filtration of these processes, i.e.,

$$\Psi_t = \text{ the } P\text{-completion of } \cap_{s>t} \sigma(z^\epsilon(y), x^\epsilon(y), u^\epsilon(y), y \leq s)$$

for each $t \geq 0$. Define a new probability measure Q on (Ω, Ψ) as before, i.e., by: $Q_t := Q|_{\Psi_t}$ is mutually absolutely continuous w.r.t. $P_t := P|_{\Psi_t}$ for all $t \geq 0$ with

$$\Lambda_t \ := \ \frac{dP_t}{dQ_t}$$

$$= exp\Big(\int_0^t \langle\gamma(z^\epsilon(s))^{-1}h(z^\epsilon(s),x^\epsilon(s),u^\epsilon(s))), d\bar{B}(s)\rangle$$

$$(3.17) \qquad - \frac{1}{2}\int_0^t \|\gamma(z^\epsilon(s))^{-1}h(z^\epsilon(s),x^\epsilon(s),u^\epsilon(s))\|^2 ds\Big),$$

where

$$\bar{B}(t) := B(t) + \int_0^t \gamma(z^\epsilon(s))^{-1}h(z^\epsilon(s),x^\epsilon(s),u^\epsilon(s))ds$$

is a d-dimensional standard Brownian motion independent of $W(\cdot)$ under Q. Furthermore,

$$\tilde{B}(t) := \int_0^t \gamma(z^\epsilon(s))^{-1}(dz^\epsilon(s) - \lambda_s(\check{h}(z^\epsilon(s),\cdot,[u_1^\epsilon(s),\cdot]))ds, \ t \geq 0,$$

is also a Brownian motion under Q adapted to $\{\mathcal{F}_t^\epsilon\}$: this again follows by Levy's martingale characterization of Brownian motion as in [14]. Since

$$\bar{B}(t) = \tilde{B}(t) + \int_0^t \gamma(z^\epsilon(s))^{-1}\lambda_s^\epsilon(\check{h}(z^\epsilon(s),\cdot,[u_1^\epsilon(s),\cdot])ds, \ t \geq 0,$$

$\bar{B}(\cdot)$ is also adapted to $\{\mathcal{F}_t^\epsilon\}$. Under Q, the dynamics (2.1)-(2.2) may be viewed in the equivalent form

$$\begin{aligned} dz^\epsilon(t) &= \gamma(z^\epsilon(z^\epsilon(t))d\bar{B}(t), \\ dx^\epsilon(t) &= \frac{1}{\epsilon}m(z^\epsilon(t),x^\epsilon(t),[u_1^\epsilon(t),u_2^\epsilon(t)])dt + \frac{1}{\sqrt{\epsilon}}\sigma(z^\epsilon(t),x^\epsilon(t))dW(t), \end{aligned}$$

where $\bar{B}(\cdot), W(\cdot)$ are independent d-, resp., s-dimensional Brownian motions.

Let $E_0[\ \cdot\ ]$ denote the expectations / conditional expectations under Q. Let $\mathcal{M}(\mathcal{R}^s)$ denote the space of finite non-negative measures on $\mathcal{R}^s$ with weak* topology. Define the 'unnormalized conditional law' as the $\mathcal{M}(\mathcal{R}^s)$-valued process $p_t^\epsilon, t \geq 0$, defined by

$$p_t^\epsilon(f) := \int f dp_t^\epsilon = E_0\left[f(x^\epsilon(t))\Lambda_t|\mathcal{F}_t^\epsilon\right]$$

$\forall$ f in a countable convergence determining class in $C_b(\mathcal{R}^s)$. From the dynamics above, it follows that for any $t \geq 0$, $(z^\epsilon(t'), u_1^\epsilon(t')), t' \geq t$, is conditionally independent of $(x^\epsilon(s), u_2^\epsilon(s)), s \leq t$, given $\mathcal{F}_t^\epsilon$. Hence for any $f \in C_0^2(\mathcal{R}^s)$,

$$(3.18) \qquad E_0\left[f(x^\epsilon(t))\Lambda_t|\mathcal{F}_{t'}^\epsilon\right] = E_0\left[f(x^\epsilon(t)\Lambda_t|\mathcal{F}_t^\epsilon\right] = p_t^\epsilon(f) \ \forall\ t' > t.$$

Then by Ito's formula, for $f \in C_0^2(\mathcal{R}^s)$, $\delta > 0$,

$$\begin{aligned} &p_{t+\delta}^\epsilon(f) - p_t^\epsilon(f) \\ &= \frac{1}{\epsilon}E_0\left[\int_t^{t+\delta} \mathcal{L}_{z^\epsilon(s)}^{u^\epsilon(s)} f(x^\epsilon(s))\Lambda_s ds|\mathcal{F}_{t+\delta}^\epsilon\right] + \\ &\quad \frac{1}{\sqrt{\epsilon}}E_0\left[\int_t^{t+\delta}\Big\langle f(x^\epsilon(s))\gamma(z^\epsilon(s))^{-1}h(z^\epsilon(s),x^\epsilon(s),u^\epsilon(s))\Lambda_s, d\bar{B}(s)\Big\rangle|\mathcal{F}_{t+\delta}^\epsilon\right]. \end{aligned}$$

Let $T > t$ and set $\delta = \frac{T-t}{N}$ for $N >> 1$. Then

$$\begin{aligned} & p^\epsilon_T(f) - p^\epsilon_t(f) \\ = & \sum_{m=0}^{N-1} (p_{t+(m+1)\delta}(f) - p_{t+m\delta}(f)) \\ = & \frac{1}{\epsilon}\sum_{m=0}^{N-1}\Big[\int_{t+m\delta}^{t+(m+1)\delta} E_0\left[\mathcal{L}^{u^\epsilon(s)}_{z^\epsilon(s)} f\Lambda_s \Big| \mathcal{F}^\epsilon_{t+(m+1)\delta}\right] ds + \\ & \frac{1}{\sqrt{\epsilon}} E_0\Big[\int_{t+m\delta}^{t+(m+1)\delta} \Big\langle \xi_s, d\bar{B}(s)\Big\rangle \Big| \mathcal{F}^\epsilon_{t+(m+1)\delta}\Big], \end{aligned}$$

where

$$\xi_s := f(x^\epsilon(s))\gamma(z^\epsilon(s))^{-1} h(z^\epsilon(s), x^\epsilon(s), u^\epsilon(s))\Lambda_s.$$

As $\delta \downarrow 0$, the first term on the right converges a.s. to

$$\frac{1}{\epsilon}\int_t^T p_s(\mathcal{L}^{u^\epsilon(s)}_{z^\epsilon(s)})ds$$

by Theorem 3.3.8, p. 56, [2] and the right continuity of σ-fields $\mathcal{F}^\epsilon_t$. We can approximate the stochastic integral

$$\int_{t+m\delta}^{t+(m+1)\delta} \Big\langle \xi_s, d\bar{B}(s)\Big\rangle$$

on the right in mean square by replacing $\{\xi_t\}$ by its approximation $\{\xi'_t\}$ with continuous paths, defined as $\xi'_t = \frac{1}{\kappa}\int_{t-\kappa}^t \xi_s ds$ for $\kappa > 0$ small, with suitable modification near $t = 0$. In turn, we can approximate the latter in mean square by replacing this $\{\xi'_t\}$ by a piecewise constant $\{\xi''_t\}$ which takes value $\xi'_{t(k)}$ on interval (say) $[t(k), t(k+1))$ where we take $\max_k |t(k+1) - t(k)|$ sufficiently small. But then the stochastic integral has the form

$$\sum_{k=k_1}^{k_2} \xi''_{t(k)}(\bar{B}(t(k+1)) - \bar{B}(t(k)))$$

where $t(k_1) = t + m\delta$ and $k_2 = t + (m+1)\delta$. But then by (3.18),

$$\begin{aligned} & E_0\left[\int_{t+m\delta}^{t+(m+1)\delta} \Big\langle \xi''_s, d\bar{B}(s)\Big\rangle | \mathcal{F}^\epsilon_{t+(m+1)\delta}\right] \\ & = \int_{t+m\delta}^{t+(m+1)\delta} \Big\langle E_0\left[\xi''_s|\mathcal{F}^\epsilon_{t+(m+1)\delta}\right], d\bar{B}(s)\Big\rangle \\ & = \int_{t+m\delta}^{t+(m+1)\delta} \Big\langle E_0\left[\xi''_s|\mathcal{F}^\epsilon_{t+m\delta}\right], d\bar{B}(s)\Big\rangle. \end{aligned}$$

By an appropriate limiting argument (in mean square) as $\delta \downarrow 0$, we get

$$p^\epsilon_T(f) - p^\epsilon_t(f) = \frac{1}{\epsilon}\int_t^T p_s(\mathcal{L}^{u^\epsilon(s)}_{z^\epsilon(s)} f)ds \ +$$

$$\frac{1}{\sqrt{\epsilon}}\int_t^T \left\langle \lambda^\epsilon_s \left(f(\cdot)\gamma^{-1}(z^\epsilon(s))\breve{h}(z^\epsilon(s),\cdot,[u^\epsilon_1(s),\cdot])\right), d\bar{B}(s)\right\rangle, \tag{3.19}$$

where we have used Theorem 3.3.8, [2] and the right continuity of sigma-fields $\{\mathcal{F}^\epsilon_t\}$ once again. This yields the evolution equation for $p^\epsilon_t, t \geq 0$:

$$\begin{aligned} dp^\epsilon_t(f) &= \frac{1}{\epsilon}p^\epsilon_t(\mathcal{L}^{u^\epsilon(t)}_{z^\epsilon(t)}f)dt \;+ \\ &\quad \frac{1}{\sqrt{\epsilon}}\langle \lambda^\epsilon_t \left(f(\cdot)\gamma^{-1}(z^\epsilon(t))\breve{h}(z^\epsilon(t),\cdot,[u^\epsilon_1(t),\cdot])\right), d\bar{B}(t)\rangle, \end{aligned} \tag{3.20}$$

for $t \geq 0$. This is the celebrated Duncan-Mortensen-Zakai equation adapted to our framework.

By the Kallianpur-Striebel abstract Bayes formula, for $f \in C_b(\mathcal{R}^s)$,

$$\pi^\epsilon_t(f) = E[f(x^\epsilon(t))|\mathcal{F}^\epsilon_t] = \frac{E_0[f(x^\epsilon(t))\Lambda_t|\mathcal{F}^\epsilon_t]}{E_0[\Lambda_t|\mathcal{F}^\epsilon_t]} = \frac{p^\epsilon_t(f)}{p^\epsilon_t(\mathbf{1})}, \tag{3.21}$$

where $\mathbf{1}$:= the constant function $\equiv 1$. Now apply Ito's formula to (3.21) to obtain (3.4). The details, though lengthy, are routine and we omit them.

Acknowledgements

This work was initiated while this author was visiting Department of Mathematics, Macquarie University. The author thanks Prof. Vladimir Gaitsgory and Prof. K. Suresh Kumar for helpful discussions. Research supported in part by a J. C. Bose Fellowship from the Government of India.

References

[1] A. Arapostathis, V. S. Borkar and M. K. Ghosh, *Ergodic Control of Diffusion Processes*, Cambridge Uni. Press, Cambridge, UK, 2012.

[2] V. S. Borkar, *Probability Theory: An Advanced Course*, Springer Verlag, New York, 1995.

[3] A. G. Bhatt and V. S. Borkar, *Occupation measures for controlled Markov processes: characterization and optimality*, The Annals of Probability **24** (1996), 1531–1562.

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

[5] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley and Sons, New York, 1986.

[6] W. H. Fleming and E. Pardoux, *Optimal control for partially observed diffusions*, SIAM J. Control and Optim. **20** (1982), 261–285.

[7] Y. Kabanov and S. Pergamenschikov, *Two-scale Stochastic Systems*, Springer-Verlag, New York, Berlin, Heidelberg, 2002.

[8] P. V. Kokotovic, H. K. Khalil and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, New York, 1986.

[9] H. J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser Boston, Boston, MA, 1990.

[10] N. I. Portenko, *Generalized Diffusion Processes*, Translations of Mathematical Monographs 83, American Math. Society, Prrovidencce, RI, 1990.

[11] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Springer Verlag, New York, 1979.

[12] K. Suresh Kumar, *Singular perturbations in stochastic ergodic control problems*, SIAM J. Control and Optim. **50** (2012), 3203–3223.
[13] D. H. Wagner, *Survey of measurable selection theorems*, SIAM. J. Control and Optim. **15** (1977), 859–903.
[14] E. Wong, *Representation of martingales, quadratic variation and applications*, SIAM J. Control **9** (1971), 621–633.

Manuscript received April 20 2016
revised October 21 2016

V. S. Borkar
Department of Electrical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India
E-mail address: borkar.vs@gmail.com