Invariance of Stochastic Control Systems with Deterministic Arguments*

Giuseppe Da Prato[†] and Hélène Frankowska[‡]

Abstract

We prove that a closed set K of a finite dimensional space is invariant under the stochastic control system

$$dX = b(X, v(t))dt + \sigma(X, v(t))dW(t), \quad v(t) \in U,$$

if and only if it is invariant under the deterministic control system with two controls

$$x' = b(x, v(t)) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x, v(t))\sigma_j(x, v(t)) + \sigma(x, v(t))u(t), \ u(t) \in H_1, \ v(t) \in U.$$

This extends the well known result of stochastic differential equations to stochastic control systems. Furthermore, we ask only $C^{1,1}$ regularity of the diffusion σ instead of the usual assumption $\sigma \in C^2$. In this way our result is new even for stochastic differential equations. The arguments of the proof are based on estimates between solutions of the stochastic control system with time independent controls and families of solutions $\{x_\omega(\cdot)\}_{\omega\in\Omega}$ to the deterministic control system

$$x' = \sigma(x, v_\omega)u_\omega(t), \quad u_\omega(t) \in H_1.$$

with appropriately chosen controls $u_\omega(t)$ and $v_\omega \in U$.

Key words. Stochastic invariance, Stratonovitch drift, stochastic control systems, state-constraints.

AMS Mathematics Subject Classification 2000 : 60H10, 93E03.

Running head : Invariance of Stochastic Control Systems

---

*Supported by the European Research Training Network “Evolution Equations for Deterministic and Stochastic Control Systems”, HPRN-CT-2002-00281.

[†]Scuola Normale Superiore di Pisa 56126, Pisa, Italy, e-mail : daprato@sns.it

[‡]CNRS, CREA, Ecole Polytechnique, 1 rue Descartes, 75005 Paris, France, e-mail : franko@poly.polytechnique.fr

1 Introduction

We are given two finite dimensional spaces H and H_1, a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ such that $\{\mathcal{F}_t\}_{t\geq 0}$ is right continuous, $\mathcal{F}_0$ contains all $\mathbf{P}-$null sets of $\mathcal{F}$ and a standard H_1-valued $\{\mathcal{F}_t\}_{t\geq 0}$-Brownian motion $W(t), t \geq 0$.

This paper is devoted to the problem of invariance of closed sets under the stochastic control system

$$dX = b(X, v(t))dt + \sigma(X, v(t))dW(t), \quad v(t) \in U, \tag{1.1}$$

where U is a complete separable metric space and $b : H \times U \to H$, and $\sigma : H \times U \to L(H_1, H)$ are bounded continuous mappings which are Lipschitz with respect to the first variable, and controls $v(t)$ are $U-$valued mappings which are progressively measurable with respect to the family $\mathcal{F}_t$, called admissible controls.

A set $K \subset H$ is invariant under the control system (1.1) if for every $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ almost surely and every admissible control $v(\cdot)$, the solution X to (1.1) starting at X_0 satisfies for all $t \geq 0$, $X(t) \in K$ almost surely. We refer to [27] for the definition of solutions to stochastic control systems.

When b and σ are control independent, the above system reduces to the stochastic differential equation

$$dX = b(X)dt + \sigma(X)dW(t). \tag{1.2}$$

Recently a number of papers were written on stochastic viability and invariance of closed sets. In the case of stochastic equation (1.2) conditions for the invariance were expressed using the Stratonovitch drift [15] (see also [17] when K is the closure of an open set with smooth boundary) or stochastic contingent sets [2, 3]. Next, a characterization of invariance, based on [15], in terms of curvature of the boundary of K was proposed in [6].

For stochastic control systems and differential inclusions different authors used stochastic contingent sets [4, 5], viscosity solutions of second order partial differential equations [8, 9, 10] and derivatives of the distance function [14], see also [18, 19, 20, 22, 23, 25] for several other approaches.

The method based on the second order partial differential equations deals with value functions of some associated optimal control problems. In [8] it is the exit time function, while in [10] it is the value function of an infinite horizon problem. These tools use the second order jets of continuous solutions to PDE's. So the second order normal cones to K arise naturally in characterizations of invariance.

In contrast, the results of [15, Doss] obtained in the context of stochastic equations use only first order normals to K. This approach is based on an equivalence between invariance of stochastic equation (1.2) and that of an associated deterministic control system. Namely it was shown in [15] that if $\sigma \in C^3$ and has bounded derivatives up to the order three, then K is invariant under the stochastic equation (1.2) if and only if K is invariant under the (well understood) deterministic control system

$$x' = b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) + \sigma(x)u(t), \quad u \in L^1_{loc}(\mathbb{R}_+, H_1). \tag{1.3}$$

Theory of [15] needs however more regularity of the diffusion term σ (C_b^3 instead of bounded and Lipschitz continuous) and is based on the support theorem which is not applicable in the presence of controls.

In this paper we prove a similar first order characterization of invariance for stochastic control systems, when $\sigma \in C_b^{1,1}$. That is we extend the Doss theorem into two directions : to control systems and less regular σ. Furthermore, we propose a very direct "deterministic proof", while in [15] arguments are based on the support theorem of stochastic analysis.

Recall that in the deterministic case a necessary and sufficient condition for invariance of K under (1.3) can be expressed by using tangents to K :

$$b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) + \sigma(x)u \in T_K(x), \ \ \forall\, x \in K, \ \forall\, u \in H_1$$

(see Section 2 for the definition of $T_K(x)$ and [1] for the thorough study of invariance in the deterministic case). The result of Doss implies that K is invariant under the stochastic system (1.2) if and only if the so called Stratonovitch drift is tangent to K:

$$b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) \in T_K(x), \ \ \forall\, x \in K \tag{1.4}$$

and the image of the diffusion σ is tangent to K :

$$\sigma(x)u \in T_K(x), \ \ \forall\, x \in K, \ \ \forall\, u \in H_1. \tag{1.5}$$

In Section 2 we show that condition (1.5) in turn is equivalent to the invariance of the boundary of K under the deterministic control system

$$x' = \sigma(x)u(t), \ \ u \in L^1_{loc}(\mathbb{R}_+, H_1). \tag{1.6}$$

Instead of using the support theorem, we take the deterministic control system (1.6) as a starting point. We first show in Section 3 that if K is invariant under (1.2), then (1.5) holds true. In fact we prove even a much stronger result for continuous data and weak solutions. Hence K is also invariant under the deterministic control system (1.6). Consider next a solution $X(t)$ to (1.2) starting at some $x \in K$. If K is invariant under (1.2), then for all $h > 0$, $X(h) \in K$ almost surely. For almost every $\omega \in \Omega$ we extend then $X_\omega(h)$ by an invariant solution of the deterministic system (1.6) with an appropriately chosen constant control u_ω. In Section 4, from an analysis of these extensions, we deduce (1.4). In this way we get two necessary conditions for the invariance (1.4) and (1.5), which are stated (equivalently) using proximal normals.

To prove that conditions (1.4) and (1.5) are also sufficient for the invariance of K under (1.2), consider a solution $X(t)$ to (1.2) starting at some random variable $X_0 \in L^2(\Omega)$ with $X_0 \in K$ almost surely. In Section 4 we check that for all $t > 0$, $\psi(t) := \mathbb{E}d_K^2(X(t)) = 0$, where $d_K(x)$ denotes the distance from x to K. The idea is to define for every fixed $t > 0$ with $\psi(t) > 0$ and all $h > 0$ an $\mathcal{F}_{t+h}-$random variable $y(h) \in K$ such that $\mathbb{E}|X(t+h) - y(h)|^2 \leq \psi(t) + Lh\psi(t) + o(h)$, with L independent from t. This leads to the inequality $d\psi(t) \leq L\psi(t)$, where $d\psi(t)$ is the lower right derivative of ψ at t. Then an extension of the Gronwall inequality proposed in Section 2 (Proposition 2.7) allows to conclude that $\psi = 0$. In order to construct $y(h)$ we use again invariant solutions to (1.6) with appropriately chosen controls.

In Section 5 we turn to stochastic control system (1.1). Taking constant controls $v \in U$, the necessary conditions for the invariance of K under (1.1) may be written as

$$b(x,v) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x,v)\sigma_j(x,v) \in T_K(x),\ \mathrm{Im}(\sigma(x,v)) \subset T_K(x),\ \forall\, v \in U,\ x \in K. \quad (1.7)$$

When controls are piecewise constant with respect to the time, by the same constructions as those used for stochastic equations in Section 4, we show that conditions (1.7) are also sufficient for the invariance. Then we approximate solutions corresponding to any admissible control by solutions with piecewise constant controls to prove the invariance of K in the full generality.

In conclusion, K is invariant under the stochastic control system (1.1) if and only if it is invariant under the deterministic control system with two (deterministic) controls

$$\begin{cases} x' = b(x,v(t)) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x,v(t))\sigma_j(x,v(t)) + \sigma(x,v(t))u(t), \\ \\ u \in L^1_{loc}(\mathbb{R}_+, H_1), \quad v : \mathbb{R}_+ \to U \ \text{ is measurable.} \end{cases}$$

In Section 5 the tangential characterization (1.7) of the invariance is also stated in terms of proximal normals and normal cones. Finally, using the same idea as in [6], but a slightly different definition, we also characterize the invariance of K under (1.1) using the curvature of K.

2 Preliminaries

We are given two euclidean finite dimensional spaces $H = \mathbb{R}^n$ and $H_1 = \mathbb{R}^m$, (norm $|\cdot|$, inner product $\langle\cdot,\cdot\rangle$) and denote by B_1 the closed unit ball in H_1, and by $\|\cdot\|$ the norm of $L(H_1, H)$.

Consider a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ such that $\{\mathcal{F}_t\}_{t\geq 0}$ is right continuous, $\mathcal{F}_0$ contains all $\mathbf{P}-$null sets of $\mathcal{F}$ and a standard H_1-valued $\{\mathcal{F}_t\}_{t\geq 0}$-Brownian motion $W(t), t \geq 0$ (see for instance [27] for the corresponding definitions). The following result is well known, since $W(t)$ is a Gaussian random variable with mean 0 and covariance operator tI.

Proposition 2.1 *There exists $C_1 > 0$ such that for all $t > 0$*

$$\mathbb{E}|W(t)|^4 \leq C_1 t^2, \ \ \mathbb{E}|W(t)|^6 \leq C_1 t^3.$$

Furthermore for any bounded adapted process $f : \mathbb{R}_+ \to L^\infty(\Omega, L(H_1, H))$ there exists $c > 0$ independent from f such that

$$\mathbb{E}\left|\int_0^t f(s)dW(s)\right|^4 \leq ct^2\|f\|^4_\infty, \ \ \forall\, t > 0.$$

Let $b : H \to H$, $\sigma : H \to L(H_1, H)$ be bounded Lipschitz continuous mappings. Denote by $\sigma^*(x)$ the transpose of $\sigma(x)$, by $\sigma_j(x)$ the column j of the matrix $\sigma(x)$ and by $D\sigma_j$ the jacobian of σ_j.

Then for every $\mathcal{F}_0$- random variable $X_0 \in L^2(\Omega)$, the differential stochastic equation

$$\begin{cases} dX = b(X)dt + \sigma(X)dW(t), \\ \\ X(0) = X_0, \end{cases} \tag{2.1}$$

has a unique strong solution $X(t)$, i.e. for all $t > 0$,

$$X(t) = X_0 + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dW(s).$$

Furthermore, for every $t_0 > 0$ there exists $M_0 > 0$ such that

$$\mathbb{E}|X(s_2) - X(s_1)|^2 \leq M_0(s_2 - s_1), \ \ \forall\, 0 \leq s_1 < s_2 \leq t_0. \tag{2.2}$$

Consider a closed non empty subset K of H. We denote by ∂K the boundary of K and by d_K the *distance* of $x \in H$ from K :

$$d_K(x) = \inf_{y \in K} |x - y|, \ x \in H.$$

Definition 2.2 *The set K is called invariant under the system* (2.1) *if for every $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ almost surely, the strong solution X to* (2.1) *satisfies for all $t \geq 0$, $X(t) \in K$ almost surely.*

Recall that the contingent cone $T_K(x)$ to K at $x \in K$ is the set of all vectors $v \in H$ such that $\liminf_{h \to 0+} d_K(x + hv)/h = 0$ and the normal cone $N_K(x)$ to K at $x \in K$ is the negative polar cone of $T_K(x)$.

Consider the set-valued map

$$\partial K \ni x \rightsquigarrow N_K(x) \subset H$$

and fix $x \in \partial K$ and $p \in N_K(x)$. The contingent derivative $DN_K(x,p)(u)$ of N_K at (x,p) in the direction $u \in H$ is defined by

$$v \in DN_K(x,p)(u) \Longleftrightarrow$$

$$\exists\, h_i \to 0+,\ \exists\, (u_i, v_i) \to (u, v) \ \text{ such that } x + h_i u_i \in \partial K, \ \ p + h_i v_i \in N_K(x + h_i u_i).$$

It is clear that $DN_K(x,p)(u) = \emptyset$, whenever $u \notin T_{\partial K}(x)$. See [7] for properties of set-valued derivatives.

The contingent curvature of K at $(x,p) \in \mathrm{Graph}(N_K)$ is defined by

$$\forall\, u,\, v \in T_{\partial K}(x), \quad \mathrm{Curv}_K(x,p)(u,v) = \sup_{\mu \in DN_K(x,p)(u)} \langle \mu, v \rangle.$$

It was introduced in [6] by the same formulae, but with the set K instead of ∂K.

Lemma 2.3 *Assume that $\sigma : H \to L(H_1, H)$ is differentiable and for every $x \in \partial K$ and $p \in N_K(x)$, $\sigma^*(x)p = 0$. Then for all $v_j \in DN_K(x,p)(\sigma_j(x))$*

$$\mathrm{Curv}_K(x,p)(\sigma_j(x), \sigma_j(x)) = \langle v_j, \sigma_j(x) \rangle = -\langle p, D\sigma_j(x)\sigma_j(x) \rangle .$$

In particular, if for every $x \in \partial K$ there exists a unique unit outward normal $n(x)$ to K at x and if $n(\cdot)$ is differentiable on ∂K, then

$$\mathrm{Tr}[n'(x)\sigma(x)\sigma(x)^*] = -\sum_{j=1}^{m} \langle n(x), D\sigma_j(x)\sigma_j(x) \rangle . \tag{2.3}$$

Proof — Let $\{e_j\}_{j=1,\dots,m}$ be an orthonormal basis of H_1, $x \in \partial K$, $p \in N_K(x)$, $u = (u^1, \dots, u^m) \in H_1$, $\mu \in DN_K(x,p)(\sigma(x)u)$ and consider $h_k \to 0+$, $\mu_k \to \mu$, $v_k \to \sigma(x)u$ such that $x + h_k v_k \in \partial K$, $p + h_k\mu_k \in N_K(x + h_k v_k)$. Then $\langle p + h_k\mu_k, \sigma(x + h_k v_k)u \rangle = 0$. Thus

$$\langle \mu_k, \sigma(x)u \rangle + \left\langle p, \sum_r \langle \nabla\sigma_{ir}(x), \sigma(x)u \rangle u^r \right\rangle = o(h_k)/h_k.$$

Taking the limit when $k \to \infty$ implies

$$\langle \mu, \sigma(x)u \rangle + \left\langle p, \sum_r \langle \nabla\sigma_{ir}(x), \sigma(x)u \rangle u^r \right\rangle = 0.$$

Setting in the last equality $u = e_j$ yields

$$\forall\, \mu \in DN_K(x,p)(\sigma_j(x)), \;\; \langle \mu, \sigma_j(x) \rangle = -\langle p, D\sigma_j(x)\sigma_j(x) \rangle . \tag{2.4}$$

To prove the last statement, observe that (2.4) yields

$$\langle \sigma(x)^* n'(x)\sigma(x)e_j, e_j \rangle = -\langle n(x), D\sigma_j(x)\sigma_j(x) \rangle .$$

Adding the above expressions for $j = 1, \dots, m$ implies

$$\mathrm{Tr}[\sigma(x)^* n'(x)\sigma(x)] = -\sum_{j=1}^{m} \langle n(x), D\sigma_j(x)\sigma_j(x) \rangle .$$

Since $\mathrm{Tr}[\sigma(x)^* n'(x)\sigma(x)] = \mathrm{Tr}[n'(x)\sigma(x)\sigma(x)^*]$ the proof is complete. □

A vector $p \in H$ is called a proximal normal to K at $x \in K$ if $|p| = d_K(x+p)$. Clearly $p = 0$ is a proximal normal and it is the only proximal normal when x is in the interior of K. It is well known that if p is a proximal normal to K at x, then for some $c > 0$

$$\forall\, y \in K, \;\; \langle p, y - x \rangle \le c|y - x|^2. \tag{2.5}$$

Proposition 2.4 *Assume that $\sigma : H \to L(H_1, H)$ is continuous. Then the following conditions are equivalent :*

(i) for all $x \in K$ and for any proximal normal p to K at x, $\sigma(x)^*p = 0$,

(ii) for all $x \in K$, $\mathrm{Im}(\sigma(x)) \subset T_K(x)$,

(iii) for all $x \in K$ and for any $p \in N_K(x)$, $\sigma(x)^*p = 0$.

Proof — Denote by $N_K^{prox}(y)$ the cone spanned by all proximal normals to K at $y \in K$. The Clarke normal cone to K at $x \in K$ is defined by

$$N_K^c(x) := \overline{co}\left(\operatorname{Limsup}_{\substack{y \to x \\ y \in K}} N_K^{prox}(y)\right),$$

where Limsup denotes the Painlevé-Kuratowski upper limit (see for instance [7]) and $\overline{co}$ the closed convex hull. The tangent cone $C_K(x)$ to K at x is the negative polar cone to $N_K^c(x)$ for $x \in K$. It is well known that $C_K(x) \subset T_K(x)$ (see [7]). From the definition of normal cone we deduce that if (i) holds true, then for all $p \in N_K^c(x)$, $\sigma(x)^\star p = 0$ implying that $\mathrm{Im}(\sigma(x)) \subset C_K(x) \subset T_K(x)$.

Observe next that if p is a proximal normal to K at x, then for every $v \in T_K(x)$, $\langle p, v\rangle \leq 0$. Consequently (ii) implies (i). Clearly (ii) yields (iii). Since for any proximal normal p to K at x we have $p \in N_K(x)$, (iii) implies (i). □

Proposition 2.5 *Assume that $\sigma : H \to L(H_1, H)$ is locally Lipschitz. Then K and ∂K are invariant under the deterministic control system*

$$y' = \sigma(y)u(t), \quad u(\cdot) \in L^1_{\mathrm{loc}}(\mathbb{R}_+, H_1) \tag{2.6}$$

if and only if and for all $x \in K$ and any proximal normal p at x we have $\sigma(x)^* p = 0$.

Proof — If K (or ∂K) is invariant under the deterministic control system (2.6), then, by taking constant controls in (2.6), from the definition of contingent cone we deduce that for all $x \in K$, $\mathrm{Im}(\sigma(x)) \subset T_K(x)$. Proposition 2.4 implies then that for all $x \in K$ and any proximal normal p at x we have $\sigma(x)^* p = 0$. Conversely, if for all $x \in K$ and any proximal normal p at x we have $\sigma(x)^* p = 0$, then by Proposition 2.4 for all $x \in K$, $\mathrm{Im}(\sigma(x)) \subset T_K(x)$. This and [1] imply that the set K is invariant under the deterministic control system (2.6).

To prove that ∂K is invariant, assume by a contradiction that for some control $u(\cdot) \in L^1_{\mathrm{loc}}(\mathbb{R}_+, H_1)$ and some $T > 0$, a solution y to (2.6) satisfies $y(0) \in \partial K$, $y(T) \in K \backslash \partial K$. Since σ is locally Lipschitz, when the control $u(\cdot)$ is fixed, solutions to (2.6) depend continuously on the initial condition. Hence there exists $y_1 \notin K$ such that the solution z to

$$z' = \sigma(z)u(t), \quad z(0) = y_1$$

satisfies $z(T) \in \mathrm{Int}(K)$. Set $x(s) = z(T-s)$. Then $x(T) = y_1 \notin K$, $x(0) = z(T) \in K$ and $x'(s) = \sigma(x(s))(-u(T-s))$. Since K is invariant under (2.6) we also have $x(T) \in K$, contradicting the choice of y_1 and completing the proof. □

Corollary 2.6 *Assume that $\sigma : H \to L(H_1, H)$ is locally Lipschitz. Then the following conditions are equivalent :*

(i) for all $x \in K$ and for any proximal normal p to K at x, $\sigma(x)^* p = 0$,

(ii) for all $x \in \partial K$, $\mathrm{Im}(\sigma(x)) \subset T_{\partial K}(x)$.

Proof — By Proposition 2.5, if (i) holds true, then ∂K is invariant under the deterministic control system (2.6). From the very definition of the contingent cone we deduce (ii). If (ii) holds true, then for all $x \in K$, $\mathrm{Im}(\sigma(x)) \subset T_K(x)$. Proposition 2.4 completes the proof. $\square$

For $\psi : \mathbb{R}_+ \to \mathbb{R}$ the lower right derivative is defined by $d\psi(t) := \liminf_{h\to 0+} \frac{\psi(t+h)-\psi(t)}{h}$.

Proposition 2.7 *Consider $T > 0$ and a continuous function $\psi : [0,T] \to \mathbb{R}_+$ with $\psi(0) = 0$. Assume that for some $L \geq 0$ and every $t \in [0,T[$ such that $\psi(t) > 0$ we have $d\psi(t) \leq L\psi(t)$. Then $\psi = 0$.*

Proof — By contradiction assume first that for some $0 < t_2 < T$, $\psi(t_2) > 0$. Let $t_0 = \max\{s \in [0,t_2] \mid \psi(s) = 0\}$. Fix any $t_\varepsilon \in (t_0,t_2)$ and set $\Psi(s) = \psi(s)$ if $s \in [t_\varepsilon,t_2]$ and $\Psi(s) = \psi(t_2)$ for all $s \geq t_2$. Then $\Psi > 0$ on $[t_\varepsilon,+\infty)$. Let K denote the epigraph of Ψ. Then K is closed and for all $(t,r) \in K$, $(1, L\Psi(t)) \in T_K(t,r)$. Hence, by the viability theorem (see [1]) the solution $(t,y(t))$ to the system

$$\begin{cases} t'(s) = 1, & t(0) = t_\varepsilon \\ y'(s) = L\Psi(t_\varepsilon + s), & y(0) = \Psi(t_\varepsilon) \end{cases}$$

satisfies $(t,y(t)) \in K$. Thus $\Psi(t_\varepsilon + t) \leq \Psi(t_\varepsilon) + \int_0^t L\Psi(t_\varepsilon + s)ds$. The Gronwall inequality implies that $\Psi(t_\varepsilon + t) \leq \Psi(t_\varepsilon)e^{Lt}$. Taking the limit when $t_\varepsilon \to t_0+$ we get $\Psi(t_0 + t) = 0$ for all $t \geq 0$. In particular $\psi(t_2) = 0$. The obtained contradiction yields the result. $\square$

3 A Necessary Condition for Viability

Consider a closed nonempty subset $K \subset H$. We first study a necessary condition for the viability of K under (1.2) and deduce from it a necessary condition for the invariance in terms of proximal normals and the diffusion. The result below may be applied to any weak solution of (1.2). See for instance [26] or [21] for the definition of weak solution.

A mapping $X : \mathbb{R}_+ \to L^2(\Omega, H)$ is called an adapted process if for every $t \geq 0$, $X(t)$ is $\mathcal{F}_t-$measurable.

Let b, σ be bounded and continuous. Assume that an adapted process $X(\cdot)$ is continuous and for some $x \in H$,

$$\forall\, t \geq 0, \;\; X(t) = x + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dW(s) \;\text{ a.s.,} \tag{3.1}$$

where a.s. states for almost surely. In the other words X is a solution to (1.2) on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ corresponding to the $\{\mathcal{F}_t\}_{t\geq 0}$-Brownian motion W and the initial condition x (in general, such solution may not exist for the given data, but it may be obtained with another probability space and Brownian motion, see [26] or [19]).

The process $X(\cdot)$ is called viable in K, if for all $t \geq 0$, $X(t) \in K$ a.s.

Theorem 3.1 *Assume that b and σ are bounded and continuous. If an adapted process $X(\cdot)$ is continuous, satisfies* (3.1) *with $x \in K$ and for some $h_i \to 0+$, $X(h_i) \in K$ a.s., then for any proximal normal p to K at x we have $\sigma(x)^* p = 0$.*

In particular, if (1.2) has a weak solution starting at some $x \in K$ which is viable in K, then for any proximal normal p to K at x we have $\sigma(x)^* p = 0$.

Corollary 3.2 *Assume that b, σ are bounded and Lipschitz continuous. If K is invariant under the system* (2.1)*, then for all $x \in K$ and any proximal normal p to K at x we have $\sigma(x)^* p = 0$.*

Proof of Corollary 3.2 By [26, Theorem 5.1.1] for every $\mathcal{F}_0-$initial condition $X_0 \in L^2(\Omega)$ there exists a continuous version of the strong solution to (2.1). Theorem 3.1 completes the proof. $\square$

Theorem 3.1 follows from a more general result below.

Theorem 3.3 *Let $f : [0,T] \to L^1(\Omega, H)$, $g : [0,T] \to L^2(\Omega, L(H_1, H))$ be adapted and continuous at zero, and $f \in L^1([0,T] \times \Omega)$, $g \in L^2([0,T] \times \Omega)$. Let $x \in K$ and define the adapted process*

$$X(t) := x + \int_0^t f(s)ds + \int_0^t g(s)dW(s).$$

Assume that at least one of the following two conditions is satisfied

(i) $X(\cdot)$ is almost surely continuous at zero,

(ii) $f(0) \in L^2(\Omega)$, $g(0) \in L^4(\Omega)$.

If for some $h_i \to 0+$, $X(h_i) \in K$ a.s., then for any proximal normal p to K at x we have $g(0)^* p = 0$ a.s.

Proof of Theorem 3.1 Consider a continuous version of X. It is enough to set $f(t) = b(X(t))$ and $g(t) = \sigma(X(t))$. Since b, σ are bounded and continuous, $f(\cdot)$ and $g(\cdot)$ are continuous at zero. All other assumptions of Theorem 3.3 are verified as well by boundedness of b, σ. $\square$

Proof of Theorem 3.3 It is not restrictive to suppose that $x = 0$ and

$$\frac{p}{|p|} = (0, ..., 1).$$

Fix $\lambda \in (0, 1/4]$. Then $d_K(\lambda p + X(h_i)) \leq \lambda |p|$, for all $i \geq 1$. Denote by $B^c(p, |p|)$ the complement of the ball $B(p, |p|)$. Set

$$\psi(y) = \left[d_{B^c(p,|p|)}(\lambda p + y) \right]^4.$$

Clearly ψ is continuous, bounded, vanishing outside of $B(p, |p|) - \lambda p$. Moreover

$$\psi(y) = \begin{cases} (|p| - |p - \lambda p - y|)^4, & \text{if } \lambda p + y \in B(p, |p|), \\ 0 & \text{otherwise.} \end{cases}$$

If $|\lambda p + y - p| < |p|$ and $p \neq \lambda p + y$, then

$$D\psi(y) = 4(|p| - |p - \lambda p - y|)^3 \, \frac{p - \lambda p - y}{|p - \lambda p - y|}.$$

If $|\lambda p + y - p| > |p|$, then $D\psi(y) = 0$. Consequently for all $y \neq p - \lambda p$,

$$D\psi(y) = \begin{cases} 4(|p| - |p - \lambda p - y|)^3 \, \dfrac{p - \lambda p - y}{|p - \lambda p - y|}, & \text{if } \lambda p + y \in B(p, |p|), \\ \\ 0 \text{ otherwise.} \end{cases}$$

So $D\psi(y)$ is bounded on $\left\{ y \in H : \lambda p + y \notin B(p, \frac{1}{4}\, |p|) \right\}$, and $D\psi(0) = 4\lambda^3 |p|^2 p$.
If $|\lambda p + y - p| < |p|$ and $p \neq \lambda p + y$, then

$$\begin{aligned} D^2\psi(y) &= 12(|p| - |p - \lambda p - y|)^2 \frac{1}{|p - \lambda p - y|^2} \, (p - \lambda p - y) \otimes (p - \lambda p - y) + \\ & \quad + 4(|p| - |p - \lambda p - y|)^3 \frac{1}{|p - \lambda p - y|^3} \, (p - \lambda p - y) \otimes (p - \lambda p - y) - \\ & \quad - 4(|p| - |p - \lambda p - y|)^3 \frac{1}{|p - \lambda p - y|} \, Id. \end{aligned}$$

If $|\lambda p + y - p| > |p|$, then $D^2\psi(y) = 0$. Consequently $D^2\psi$ is continuous and bounded outside of $B(p - \lambda p, \frac{1}{4}\ |p|)$. Differentiating again we find that $D^3\psi$ is continuous and bounded on the set

$$\left\{ y \in H : \lambda p + y \notin B(p, \frac{1}{4}\ |p|) \right\}.$$

Moreover

$$D^2\psi(0) = 12\lambda^2 p \otimes p + \frac{4\lambda^3}{1 - \lambda}\, p \otimes p - 4|p|^2 \frac{\lambda^3}{1 - \lambda}\, Id.$$

Let $\rho \in C^\infty$ be such that $0 \le \rho \le 1$ and

$$\rho(y) = \begin{cases} 0, & \text{if } \lambda p + y \in B(p, \frac{1}{4}\ |p|), \\ \\ 1, & \text{if } \lambda p + y \in B^c(p, \frac{1}{2}\ |p|). \end{cases}$$

Define the function $\zeta(y) = \rho(y)\psi(y)$. Then $\zeta \in C^2$, is bounded and has bounded derivatives up to order 3. Furthermore,

$$\zeta(y) \le \lambda^4 |p|^4, \quad \forall\, y \in K. \tag{3.2}$$

From Itô's formula (see for instance [21, Theorem 4.2.1]) and (3.2), since $X(h_i) \in K$ a.s. we get

$$\begin{aligned} \zeta(X(h_i)) &= \lambda^4 |p|^4 + \int_0^{h_i} \left(\frac{1}{2} \operatorname{Tr} \left[D^2\zeta(X(s)) g(s) g(s)^* \right] + \langle D\zeta(X(s)), f(s) \rangle \right) ds \\ & \quad + \int_0^{h_i} \langle D\zeta(X(s)), g(s) dW(s) \rangle \le \lambda^4 |p|^4 \ \text{ a.s.} \end{aligned}$$

Taking the expectation and using Fubini's theorem to bring the expectation inside the integrals, we obtain

$$\int_0^{h_i} \mathbb{E}\langle D\zeta(X(s)), f(s)\rangle ds + \frac{1}{2}\int_0^{h_i} \mathbb{E}\ \mathrm{Tr}\ [D^2\zeta(X(s))g(s)g(s)^*]ds \leq 0. \tag{3.3}$$

Since $D\zeta$ is bounded, for a constant c_1 independent of s

$$\langle D\zeta(0), f(0)\rangle \leq \langle D\zeta(X(s)), f(s)\rangle + c_1|f(s) - f(0)| + |D\zeta(X(s)) - D\zeta(0)||f(0)|. \tag{3.4}$$

If the assumption (i) holds true, then $|D\zeta(X(s)) - D\zeta(0)||f(0)| \to 0$ a.s. when $s \to 0+$. This and the dominated convergence theorem yield

$$\int_0^{h_i} \mathbb{E}|D\zeta(X(s)) - D\zeta(0)||f(0)|ds = o(h_i). \tag{3.5}$$

If (ii) is satisfied, then, by the Lipschitz continuity of $D\zeta$, for some $L > 0$,

$$|D\zeta(X(s)) - D\zeta(0)||f(0)| \leq L|X(s)||f(0)|$$

and, by the Höder inequality and (2.2), for some $c_2 > 0$ and all i,

$$\begin{aligned}\int_0^{h_i} \mathbb{E}|D\zeta(X(s)) - D\zeta(0)||f(0)|ds &\leq L\int_0^{h_i} \|X(s)\|_{L^2}\|f(0)\|_{L^2}ds \\ &\leq c_2L\|f(0)\|_{L^2}\int_0^{h_i} \sqrt{s}ds = o(h_i).\end{aligned}$$

Thus also in this case we have (3.5).

Taking the expectation in (3.4) and integrating on $[0, h_i]$, we deduce from (3.5) and the continuity of f at zero that

$$h_i\,\mathbb{E}\langle D\zeta(0), f(0)\rangle \leq \int_0^{h_i} \mathbb{E}\langle D\zeta(X(s)), f(s)\rangle ds + o(h_i). \tag{3.6}$$

On the other hand, since $D^2\zeta$ is bounded, for some $c_3 > 0$ independent from s,

$$\begin{aligned}&\mathrm{Tr}\ [D^2\zeta(0)g(0)g(0)^*] \leq\ \mathrm{Tr}\ [D^2\zeta(X(s))g(s)g(s)^*]+ \\ &+\|D^2\zeta(X(s)) - D^2\zeta(0)\|\|g(0)\|^2 + c_3\|g(s) - g(0)\|\left(\|g(s)\| + \|g(0)\|\right).\end{aligned} \tag{3.7}$$

If (i) holds true, then $\|D^2\zeta(X(s)) - D\zeta(0)\|\|g(0)\|^2 \to 0$ a.s. when $s \to 0+$. This and the dominated convergence theorem yield

$$\int_0^{h_i} \mathbb{E}\|D^2\zeta(X(s)) - D^2\zeta(0)\|\|g(0)\|^2ds = o(h_i). \tag{3.8}$$

If (ii) is satisfied, then, by the Lipschitz continuity of $D^2\zeta$, for some $L > 0$,

$$\|D^2\zeta(X(s)) - D^2\zeta(0)\|\|g(0)\|^2 \leq L|X(s)|\|g(0)\|^2$$

and, by the Hölder inequality, for some $c_4 > 0$ independent from s

$$\mathbb{E}\left(|X(s)|\|g(0)\|^2\right) \leq \|X(s)\|_{L^2}\|g(0)\|^2_{L^4} \leq c_4\|g(0)\|^2_{L^4}\sqrt{s}.$$

Thus also in this case we have (3.8). On the other hand, by the Hölder inequality,

$$\mathbb{E}\left(\|g(s)\|\|g(s)-g(0)\|\right) \leq \|g(s)\|_{L^2}\|g(s)-g(0)\|_{L^2},$$

and

$$\mathbb{E}\left(\|g(0)\|\|g(s)-g(0)\|\right) \leq \|g(0)\|_{L^2}\|g(s)-g(0)\|_{L^2}.$$

Taking the expectation in (3.7) and integrating on $[0, h_i]$, we deduce from (3.8), the continuity of g at zero and the last two inequalities that

$$\frac{h_i}{2}\,\mathbb{E}\mathrm{Tr}\left[D^2\zeta(0)g(0)g(0)^*\right] \leq \frac{1}{2}\int_0^{h_i}\mathbb{E}\mathrm{Tr}\left[D^2\zeta(X(s))g(s)g(s)^*\right]ds + o(h_i). \tag{3.9}$$

Consequently, by (3.3), (3.6), (3.9)

$$h_i\,\mathbb{E}\left(\langle D\zeta(0), f(0)\rangle + \frac{1}{2}\,\mathrm{Tr}\left[D^2\zeta(0)g(0)g(0)^*\right]\right) \leq o(h_i).$$

Dividing by h_i and taking the limit, we obtain

$$\mathbb{E}\langle D\zeta(0), f(0)\rangle + \frac{1}{2}\,\mathbb{E}\mathrm{Tr}\,\left[D^2\zeta(0)g(0)g(0)^*\right] \leq 0.$$

Thus

$$4\lambda^3|p|^2\mathbb{E}\langle p, f(0)\rangle + 6\lambda^2\,\mathbb{E}\mathrm{Tr}\left[(p\otimes p)g(0)g(0)^*\right]+$$

$$+\lambda^3\left[\frac{2}{1-\lambda}\,\mathbb{E}\mathrm{Tr}\left[(p\otimes p)g(0)g(0)^*\right] - \frac{2}{1-\lambda}\,|p|^2\mathbb{E}\mathrm{Tr}\left[g(0)g(0)^*\right]\right] \leq 0.$$

Dividing by λ^2 and letting $\lambda \to 0$ we deduce that

$$\mathbb{E}\mathrm{Tr}\left[(p\otimes p)g(0)g(0)^*\right] \leq 0.$$

But $p\otimes p = (\alpha_{i,j})$ with $\alpha_{n,n} = 1$ and $\alpha_{i,j} = 0$ for $i + j < 2n$. Thus

$$\mathbb{E}\mathrm{Tr}\left[(p\otimes p)g(0)g(0)^*\right] = \mathbb{E}\sum_{i=1}^{n} g_{n,i}(0)^2 \leq 0.$$

So $g_{n,i}(0) = 0$ a.s. for all i and $g(0)^*p = 0$ a.s. as required. $\Box$

Remark 3.4 In [2] the authors investigate invariance by using stochastic tangent sets. Namely a pair of $\mathcal{F}_t$-random variables $(\beta,\gamma) \in L^2(\Omega, H) \times L^2(\Omega, L(H_1, H))$ is tangent to K at $(t, x) \in \mathbb{R}_+ \times K$ if there exist continuous adapted processes $\zeta(s)$ and $\eta(s)$ converging to zero when $s \to t+$ such that for all small $h > 0$,

$$x + \int_t^{t+h}(\beta + \zeta(s))ds + \int_t^{t+h}(\gamma + \eta(s))dW(s) \in K \text{ a.s.}$$

If for some $T > t$, $\zeta \in L^1([t, T]\times\Omega)$, $\eta \in L^2([t, T]\times\Omega)$ and $\gamma \in L^4(\Omega)$, then Theorem 3.3 may be applied and thus we get $\gamma^*p = 0$ a.s. for every proximal normal p to K at x.

4 Necessary and Sufficient Conditions for the Invariance

We denote by $C_b^{1,1}(H; L(H_1, H))$ the set of all functions $\sigma : H \to L(H_1, H)$ such that σ and its derivative σ' are bounded and $\sigma'(\cdot)$ is Lipschitz.

In this section we assume that K is closed, b is Lipschitz and bounded, and that $\sigma \in C_b^{1,1}(H; L(H_1, H))$. Recall that $\sigma_j(x)$ denotes the column j of the matrix $\sigma(x)$ and $D\sigma_j$ the jacobian of σ_j.

Theorem 4.1 *Assume that K is closed, b is Lipschitz and bounded, and that $\sigma \in C_b^{1,1}$. The set K is invariant under the system* (2.1) *if and only if for every $x \in \partial K$ and for all proximal normal p to K at x we have*

$$\left\langle p,\ b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) \right\rangle \leq 0, \quad \sigma(x)^* p = 0. \tag{4.1}$$

Remark 4.2

a) In Theorem 4.1 instead of taking proximal normals p, we may take vectors $p \in N_K^c(x)$ (the Clarke normal cone to K at x). Indeed, it is enough to apply the same arguments as those from the proof of Proposition 2.4 to show that if (4.1) holds true for all proximal normals, then it is also valid for all $p \in N_K^c(x)$. Conversely, since $N_K^{prox}(x) \subset N_K^c(x)$, if (4.1) holds true for all $p \in N_K^c(x)$, then it is also valid for all proximal normals to K at x.

b) In [8] invariance is defined in a different way. Namely the initial conditions are elements of H instead of random variables.

From the proof of Theorem 4.1 provided below it follows that (4.1) is also a necessary and sufficient condition for the invariance of K with this different definition.

Since $N_K^{prox}(x) \subset N_K(x) \subset N_K^c(x)$ the above remark implies the following corollary.

Corollary 4.3 *Assume that K is closed, b is Lipschitz and bounded, and that $\sigma \in C_b^{1,1}$. Then K is invariant under* (2.1) *if and only if for every $x \in \partial K$ and for all $p \in N_K(x)$ the relations* (4.1) *hold true.*

Theorem 4.1 allows to extend to a less regular σ a result from [15] :

Corollary 4.4 *Assume that K is closed, b is Lipschitz and bounded, and that $\sigma \in C_b^{1,1}$. Then K is invariant under the stochastic system* (2.1) *if and only if K is invariant under the deterministic control system*

$$x' = b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) + \sigma(x)u(t), \quad u \in L^1_{loc}(\mathbb{R}_+, H_1), \tag{4.2}$$

or, equivalently, if and only if

$$b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) \in T_K(x), \quad \mathrm{Im}(\sigma(x)) \subset T_K(x), \ \ \forall\, x \in K.$$

Proof — By Corollary 4.3, relations (4.1) may be equivalently written with $p \in N_K(x)$ instead of proximal normals. Thus, by the separation theorem, (4.1) is equivalent to

$$\forall\, x \in K,\ \ b(x) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x)\sigma_j(x) \in \overline{\mathrm{co}}\, T_K(x)\ \ \&\ \ \mathrm{Im}(\sigma(x)) \subset \overline{\mathrm{co}}\, T_K(x),$$

where $\overline{\mathrm{co}}$ denotes the closed convex hull. By [7, Theorem 4.1.10] and continuity of b, σ in the above $\overline{\mathrm{co}}\, T_K(x)$ may be replaced by $T_K(x)$. Hence, by [1], (4.1) is a necessary and sufficient condition for the invariance of K under (4.2). $\square$

Theorem 4.1 and Lemma 2.3 imply the following two corollaries.

Corollary 4.5 *Assume that b is Lipschitz and bounded, that $\sigma \in C_b^{1,1}$, K is closed and for every $x \in \partial K$ there exists a unique unit outward normal $n(x)$ to K at x. If $n(\cdot)$ is differentiable on ∂K, then K is invariant under* (2.1) *if and only if for every $x \in \partial K$*

$$\langle n(x), b(x)\rangle + \frac{1}{2}\mathrm{Tr}[n'(x)\sigma(x)\sigma^*(x)] \le 0,\ \ \sigma(x)^* n(x) = 0. \tag{4.3}$$

Corollary 4.6 *Assume that K is closed, b is Lipschitz and bounded and that $\sigma \in C_b^{1,1}$. Further assume that for every $x \in \partial K$ and $p \in N_K(x)$, $DN_K(x,p)(\sigma_j(x)) \neq \emptyset$ for all $j = 1, ..., m$. Then K is invariant under* (2.1) *if and only if for every $x \in \partial K$ and $p \in N_K(x)$*

$$\langle p, b(x)\rangle + \frac{1}{2}\sum_{j=1}^{m} \mathrm{Curv}_K(x,p)(\sigma_j(x), \sigma_j(x)) \le 0,\ \ \sigma(x)^* p = 0. \tag{4.4}$$

Furthermore, for all $v_j \in DN_K(x,p)(\sigma_j(x))$

$$\mathrm{Curv}_K(x,p)(\sigma_j(x), \sigma_j(x)) = \langle v_j, \sigma_j(x)\rangle = -\langle p, D\sigma_j(x)\sigma_j(x)\rangle\,.$$

In order to prove Theorem 4.1 we shall need the following lemma. Let $x \in H$ and let $X(\cdot)$ be the strong solution to (2.1) with $X_0 = x$. Set

$$\begin{aligned} I_{ij}(t) \ &:= \sigma_{ij}(X(t)) - \sigma_{ij}(x) - \\ &- \int_0^t \langle \nabla\sigma_{ij}(X(s)), b(X(s))\rangle\, ds - \int_0^t \langle \nabla\sigma_{ij}(X(s)), \sigma(X(s))dW(s)\rangle\,. \end{aligned} \tag{4.5}$$

Lemma 4.7 *There exists a constant $M_1 > 0$ independent from x such that for all $1 \le i \le n$, $1 \le j \le m$*

$$\mathbb{E}(I_{ij}(t))^2 \le M_1 t^2,\ \ \mathbb{E}(I_{ij}(t))^4 \le M_1 t^4,\ \ \forall\, t \ge 0. \tag{4.6}$$

Proof — We first assume that in addition $\sigma \in C^2$. By the Itô formula

$$\begin{aligned} \sigma_{ij}(X(t)) \ &= \sigma_{ij}(x) + \int_0^t \left(\langle \nabla\sigma_{ij}(X(s)), b(X(s))\rangle + \frac{1}{2}\mathrm{Tr}[\sigma_{ij}''\sigma\sigma^*](X(s)) \right) ds + \\ &+ \int_0^t \langle \nabla\sigma_{ij}(X(s)), \sigma(X(s))dW(s)\rangle\,. \end{aligned}$$

Thus $\mathbb{E}(I_{ij}(t))^r = \mathbb{E}\left(\frac{1}{2}\int_0^t \mathrm{Tr}[\sigma''_{ij}\sigma\sigma^*](X(s))ds\right)^r$ for $r = 2,\ 4$. Since σ''_{ij} and σ are bounded, it is enough to take $M_1 = n^4\left(1 + \max_{i,j}\|\sigma''_{ij}\sigma\sigma^*\|_\infty\right)^4$, where n =dim(H).

Consider next a mollifier $\psi : H \to [0, 1]$ and the sequence of C^∞ functions $\{\sigma^k_{ij}\}_{k\geq 1}$ defined by

$$\sigma^k_{ij}(z) := \int \sigma_{ij}(z - y/k)\psi(y)dy.$$

Then σ^k_{ij} and its first and second derivatives are bounded with the same bounds than those of σ. Furthermore, $\sigma^k_{ij},\ \nabla\sigma^k_{ij}$ converge pointwise to σ_{ij} and $\nabla\sigma_{ij}$ respectively. Define $I^k_{ij}(t)$ as in (4.5) with σ_{ij} replaced by σ^k_{ij}. By the first part of the proof for some $c > 0$ and all $k \geq 1$, $\mathbb{E}(I^k_{ij}(t))^2 \leq ct^2$. Set

$$J_k := \int_0^t \left\langle \left(\sigma^*\nabla\sigma_{ij} - (\sigma^k)^*\nabla\sigma^k_{ij}\right)(X(s)), dW(s)\right\rangle.$$

Then for some constant $\alpha > 1$

$$\frac{1}{\alpha}\mathbb{E}(I_{ij}(t))^2 \leq \mathbb{E}(I^k_{ij}(t))^2 + \mathbb{E}\left(\sigma_{ij}(X(t)) - \sigma^k_{ij}(X(t))\right)^2 + (\sigma_{ij}(x) - \sigma^k_{ij}(x))^2 +$$

$$+\mathbb{E}\left(\int_0^t \left\langle(\nabla\sigma_{ij} - \nabla\sigma^k_{ij})(X(s)), b(X(s))\right\rangle ds\right)^2 + \mathbb{E}(J_k)^2 \leq$$

$$\leq ct^2 + \mathbb{E}\left(\sigma_{ij}(X(t)) - \sigma^k_{ij}(X(t))\right)^2 + (\sigma_{ij}(x) - \sigma^k_{ij}(x))^2 +$$

$$+t||b||^2_\infty \int_0^t \mathbb{E}|(\nabla\sigma_{ij} - \nabla\sigma^k_{ij})(X(s))|^2 ds + \int_0^t \mathbb{E}|(\sigma^*\nabla\sigma_{ij} - (\sigma^k)^*\nabla\sigma^k_{ij})(X(s))|^2 ds.$$

Taking the limit when $k \to \infty$ we end the proof of the first inequality in (4.6) with the constant $M_1 = \alpha c$.

To prove the second inequality, observe that for some constant $\beta > 1$

$$\frac{1}{\beta}\mathbb{E}(I_{ij}(t))^4 \leq \mathbb{E}(I^k_{ij}(t))^4 + \mathbb{E}\left(\sigma_{ij}(X(t)) - \sigma^k_{ij}(X(t))\right)^4 + (\sigma_{ij}(x) - \sigma^k_{ij}(x))^4 +$$

$$+\mathbb{E}\left(\int_0^t \left\langle(\nabla\sigma_{ij} - \nabla\sigma^k_{ij})(X(s)), b(X(s))\right\rangle ds\right)^4 + \mathbb{E}(J_k)^4.$$

We may apply the same limiting argument provided we show that

$$\lim_{k\to\infty} \mathbb{E}|J_k|^4 = 0.$$

By Proposition 2.1 for a constant $c_0 > 0$

$$\mathbb{E}|J_k|^4 \leq c_0 t^2 \|\sigma^*\nabla\sigma_{ij} - (\sigma^k)^*\nabla\sigma^k_{ij}\|^4_\infty.$$

The right-hand side of the above inequality converging to zero, the proof follows. $\square$

Fix $u = (u^1, ..., u^m) \in B_1$ and consider the ordinary differential equation

$$y' = \sigma(y)u, \quad y(0) = x.$$

Since $\sigma \in C_b^{1,1}$, its solution $y(\cdot)$ verifies

$$y(s) = x + s\sigma(x)u + \frac{s^2}{2}\left(\sum_j \langle \nabla\sigma_{ij}(x), \sigma(x)u\rangle\, u^j\right) + O_x(s^3), \tag{4.7}$$

where for some $M > 0$ independent from x, and all $s \geq 0$

$$|O_x(s^3)| \leq Ms^3. \tag{4.8}$$

Proof of Theorem 4.1 —

Necessity. Assume that K is invariant. Then, by Corollary 3.2, $\sigma^*(y)p_y = 0$ for all $y \in K$ and any proximal normal p_y to K at y.

Fix $x \in K$ and a proximal normal p to K at x. Then for some $c > 0$ we have

$$\forall\, z \in K, \quad \langle p, z - x\rangle \leq c|z - x|^2. \tag{4.9}$$

Consider the strong solution X to (2.1) with $X_0 = x$. Then for a constant $M_0 > 0$ and $t_0 = 1$ the inequality (2.2) holds true. Fix $0 < t \leq 1$, and also an element $\widetilde{W}(t)$ in the class of functions equivalent to $W(t) \in L^2(\Omega; H_1)$, and an element $\widetilde{X}(t)$ in the class of functions equivalent to $X(t) \in L^2(\Omega; H)$. For every $\omega \in \Omega$ define

$$u_\omega := \begin{cases} 0 & \text{if } \widetilde{W}_\omega(t) = 0, \\ -\dfrac{\widetilde{W}_\omega(t)}{|\widetilde{W}_\omega(t)|} & \text{otherwise.} \end{cases}$$

For all $\omega \in \Omega$, let $z_\omega(\cdot)$ be the solution to the deterministic system

$$z'(s) = \sigma(z(s))u_\omega, \quad z(0) = \widetilde{X}_\omega(t). \tag{4.10}$$

Since $\widetilde{X}(t) \in K$ almost surely, by Proposition 2.5 we know that for almost all $\omega \in \Omega$ and for all $s \geq 0$, $z_\omega(s) \in K$.

For all $\xi \in H$ define $F(t, \xi) : \Omega \to L(H_1, H)$ by

$$F(t, \xi) := (f_{ij}(t, \xi)), \quad f_{ij}(t, \xi)_\omega = \langle \nabla\sigma_{ij}(\xi), \sigma(\xi)\widetilde{W}_\omega(t)\rangle.$$

By (4.7)

$$z_\omega(|\widetilde{W}_\omega(t)|) = \widetilde{X}_\omega(t) - \sigma(\widetilde{X}_\omega(t))\widetilde{W}_\omega(t) + \frac{1}{2}F(t, \widetilde{X}_\omega(t))\widetilde{W}_\omega(t) + \Phi_\omega(t), \tag{4.11}$$

where $|\Phi_\omega(t)| \leq M|\widetilde{W}_\omega(t)|^3$. Set $y_\omega(t) = z_\omega(|\widetilde{W}_\omega(t)|)$. We claim that $y(t)$ is $\mathcal{F}_t-$measurable and $y(t) \in L^2(\Omega, H)$. Indeed for all $(x_0, v_0) \in H \times H_1$ consider the solution $z(\cdot; x_0, v_0)$ to $z' = \sigma(z)v_0$ satisfying $z(0) = x_0$ and define the closed set

$$\Pi := \{(s, x_0, v_0, z(s; x_0, v_0)) \mid s \geq 0,\ x_0 \in H,\ v_0 \in H_1\}.$$

Observe that the set $\Gamma(t) := \{(\omega, |\widetilde{W}_\omega(t)|, \widetilde{X}_\omega(t), u_\omega, r) \mid \omega \in \Omega,\ r \in H\}$ belongs to $\mathcal{F}_t \times \mathcal{B}_1 \times \mathcal{B}_H \times \mathcal{B}_{H_1} \times \mathcal{B}_H$, where $\mathcal{B}_1$ denotes the $\sigma-$ algebra of Borel subsets of $\mathbb{R}$, $\mathcal{B}_H$ denotes the $\sigma-$ algebra of Borel subsets of H and $\mathcal{B}_{H_1}$ denotes the $\sigma-$ algebra of Borel subsets of H_1. Then $\Gamma(t) \cap (\Omega \times \Pi) \in \mathcal{F}_t \times \mathcal{B}_1 \times \mathcal{B}_H \times \mathcal{B}_{H_1} \times \mathcal{B}_H$. Since

$$\Gamma(t) \cap (\Omega \times \Pi) = \{(\omega, |\widetilde{W}_\omega(t)|, \widetilde{X}_\omega(t), u_\omega, y_\omega(t)) \mid \omega \in \Omega\},$$

by the projection theorem (see for instance [11]) $\omega \mapsto y_\omega(t)$ is $\mathcal{F}_t-$measurable. Since σ is bounded, we deduce that $y(t) \in L^2(\Omega, H)$. Define $G(t) : \Omega \to L(H_1, H)$ by

$$G(t) = (g_{ij}(t)), \quad g_{ij}(t) = \int_0^t \langle \nabla\sigma_{ij}(X(s)), \sigma(X(s))dW(s)\rangle.$$

Applying Lemma 4.7 to (4.11) we obtain

$$\begin{aligned} y(t) \;\; = \;\; & x + \int_0^t b(X(s))ds - \left(\int_0^t \langle \nabla\sigma_{ij}(X(s)), b(X(s))\rangle ds + G(t)\right)\widetilde{W}(t) + \\ & + \int_0^t (\sigma(X(s)) - \sigma(x))dW(s) + \frac{1}{2}F(t, \widetilde{X}(t))\widetilde{W}(t) + \Phi(t) + \Psi(t)\widetilde{W}(t), \end{aligned} \tag{4.12}$$

where $|\Phi_\omega(t)| \leq M|\widetilde{W}_\omega(t)|^3$ and $\mathbb{E}\|\Psi(t)\|^2 \leq Ct^2$, $\mathbb{E}\|\Psi(t)\|^4 \leq Ct^4$ for some constant $C > 0$ independent from t. Since $y(t) \in K$ almost surely, by (4.9),

$$\mathbb{E}\langle p, y(t) - x\rangle \leq c\,\mathbb{E}|y(t) - x|^2. \tag{4.13}$$

By (4.12) for some $\alpha > 1$

$$\begin{aligned} & \frac{1}{\alpha}|y(t) - x|^2 \leq t^2\,\|b\|_\infty + t^2|\widetilde{W}(t)|^2 \max_{i,j} \|\langle \nabla\sigma_{ij}, b\rangle\|_\infty^2 + \\ & + |\widetilde{W}(t)|^2 \Sigma_{i,j} \left(\int_0^t \langle \nabla\sigma_{ij}(X(s)), \sigma(X(s))dW(s)\rangle\right)^2 + \\ & + \left|\int_0^t (\sigma(X(s)) - \sigma(x))dW(s)\right|^2 + |\widetilde{W}(t)|^4 \max_{i,j} \|\sigma^*\nabla\sigma_{ij}\|_\infty^2 + |\Phi(t)|^2 + \\ & + \|\Psi(t)\|^2|\widetilde{W}(t)|^2 = I_1(t) + I_2(t) + I_3(t) + I_4(t) + I_5(t) + I_6(t) + I_7(t). \end{aligned} \tag{4.14}$$

Then $\mathbb{E}(I_1(t)) = O(t^2)$ and since $\mathbb{E}|W(t)|^2 = t$, $\mathbb{E}(I_2(t)) = O(t^3)$. By Proposition 2.1 for some $c_1 > 0$ independent from $t \in (0, 1]$

$$(\mathbb{E}(I_3(t)))^2 \leq c_1 \mathbb{E}|W(t)|^4 \ \max_{i,j} \mathbb{E}\left(\int_0^t \langle \nabla\sigma_{ij}(X(s)), \sigma(X(s))dW(s)\rangle\right)^4 = O(t^4).$$

Consequently, $\mathbb{E}(I_3(t)) = O(t^2)$. By the Lipschitz continuity of σ and (2.2)

$$\mathbb{E}(I_4(t)) = \int_0^t \mathbb{E}|\sigma(X(s)) - \sigma(x)|^2 ds = O(t^2).$$

By Proposition 2.1, $\mathbb{E}(I_5(t)) = O(t^2)$. We also know that $\mathbb{E}|\Phi(t)|^2 \le M\mathbb{E}|\int_0^t dW(s)|^6$. By Proposition 2.1,

$$\mathbb{E}|\Phi(t)|^2 = O(t^3). \tag{4.15}$$

Thus, $\mathbb{E}(I_6(t)) = O(t^3)$. Finally observe that

$$\mathbb{E}(I_7(t)) \le \mathbb{E}\|\Psi(t)\|^4 + \mathbb{E}|W(t)|^4 = O(t^2).$$

The above estimates and (4.14) imply that for all $0 \le t \le 1$, $\mathbb{E}|y(t) - x|^2 = O(t^2)$ and, by (4.13),

$$\mathbb{E}\left\langle p, y(t) - x\right\rangle = O(t^2). \tag{4.16}$$

By (4.12) we have

$$\begin{aligned}
&\mathbb{E}\langle p, tb(x)\rangle - \mathbb{E}\left\langle p, G(t)W(t)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle p, F(t, \widetilde{X}(t))W(t)\right\rangle = \\
&= \mathbb{E}\langle p, y(t) - x\rangle - \mathbb{E}\langle p, \int_0^t (b(X(s)) - b(x))ds\rangle + \\
&+\mathbb{E}\left\langle p, \left(\int_0^t \langle\nabla\sigma_{ij}(X(s)), b(X(s))\rangle ds\right) W(t)\right\rangle - \mathbb{E}\langle p, \Phi(t) + \Psi(t)W(t)\rangle = \\
&= A_1(t) + A_2(t) + A_3(t) + A_4(t).
\end{aligned} \tag{4.17}$$

By the Lipschitz continuity of b, Fubini's theorem, the Hölder inequality and (2.2) for some $c_2 > 0$,

$$A_2(t) \le c_2 \int_0^t \sqrt{s}ds \le c_2 t^{3/2}.$$

Furthermore, for some $c_3 > 0$

$$A_3(t) \le c_3 t\sqrt{\mathbb{E}|W(t)|^2} = c_3 t^{3/2}.$$

Moreover

$$\mathbb{E}\langle -p, \Psi(t)W(t)\rangle \le |p|\sqrt{\mathbb{E}\|\Psi(t)\|^2 \mathbb{E}|W(t)|^2} \le |p|\sqrt{Ct^3}$$

and, by (4.15),

$$\mathbb{E}\langle -p, \Phi(t)\rangle \le |p|\sqrt{\mathbb{E}\Phi(t)^2} = O(t^{3/2)}.$$

Hence from (4.16) and (4.17) we deduce that for all $0 < t \le 1$

$$\mathbb{E}\langle p, tb(x)\rangle - \mathbb{E}\left\langle p, G(t)W(t)\right\rangle + \frac{1}{2}\mathbb{E}\left\langle p, F(t, \widetilde{X}(t))W(t)\right\rangle = O(t^{3/2}). \tag{4.18}$$

By the Lipschitz continuity of σ and $\nabla\sigma_{ij}$ and by (2.2), for some $c_4 > 0$ and all $s \in [0, 1]$

$$\mathbb{E}|\sigma^*(X(s))\nabla\sigma_{ij}(X(s)) - \sigma^*(x)\nabla\sigma_{ij}(x)|^2 \le c_4 s. \tag{4.19}$$

This and the Hölder inequality imply that for all $1 \leq i \leq n,\ 1 \leq j \leq m$

$$\mathbb{E}|\langle \nabla\sigma_{ij}(X(t)), \sigma(X(t))W(t)\rangle W^j(t) - \langle \nabla\sigma_{ij}(x), \sigma(x)W(t)\rangle W^j(t)| \leq$$

$$\leq \sqrt{c_4 t}\sqrt{\mathbb{E}|W(t)|^4} = O(t^{3/2}).$$

Consequently,

$$\mathbb{E}|F(t, \widetilde{X}(t))W(t) - F(t,x)W(t)| = O(t^{3/2}). \tag{4.20}$$

Furthermore, by the Hölder inequality and (4.19), for a constant $c_5 > 0$

$$\mathbb{E}\left|\left\langle p, \left(\int_0^t \langle \sigma(X(s))^* \nabla\sigma_{ij}(X(s)) - \sigma(x)^*\nabla\sigma_{ij}(x), dW(s)\rangle\right) W(t)\right\rangle\right| \leq$$

$$\leq c_5\sqrt{t}\ \max_{i,j}\left(\mathbb{E}|\int_0^t \langle \sigma(X(s))^*\nabla\sigma_{ij}(X(s)) - \sigma(x)^*\nabla\sigma_{ij}(x), dW(s)\rangle|^2\right)^{1/2} = \tag{4.21}$$

$$= c_5\sqrt{t}\ \max_{i,j}\left(\int_0^t \mathbb{E}|\sigma^*(X(s))\nabla\sigma_{ij}(X(s)) - \sigma^*(x)\nabla\sigma(x)|^2 ds\right)^{1/2} = O(t^{3/2}).$$

Thus it follows from (4.18), (4.20) and (4.21) that for all $0 < t \leq 1$

$$\begin{aligned} &\mathbb{E}\langle p, tb(x)\rangle + \frac{1}{2}\mathbb{E}\,\langle p, F(t,x)W(t)\rangle - \\ &-\mathbb{E}\left\langle p, \left(\int_0^t \langle \nabla\sigma_{ij}(x), \sigma(x)dW(s)\rangle\right) W(t)\right\rangle = O(t^{3/2}). \end{aligned} \tag{4.22}$$

Finally observe that

$$\begin{aligned} &\mathbb{E}\left\langle p, \left(\int_0^t \langle \nabla\sigma_{ij}(x), \sigma(x)dW(s)\rangle\right) W(t)\right\rangle = \\ &= \mathbb{E}\left\langle p, \left(\sum_j \langle \nabla\sigma_{ij}(x), \sigma(x)W(t)\rangle W^j(t)\right)\right\rangle = \mathbb{E}\langle p, F(t,x)W(t)\rangle = \\ &= \mathbb{E}\left(\sum_i p_i \sum_{j,k,r} \frac{\partial\sigma_{ij}}{\partial x_k}(x)\sigma_{kr}(x)W^r(t)W^j(t)\right) = \\ &= \sum_i p_i \sum_{j,k} \frac{\partial\sigma_{ij}}{\partial x_k}(x)\sigma_{kj}(x)\mathbb{E}(W^j(t))^2 = t\left\langle p, \sum_j D\sigma_j(x)\sigma_j(x)\right\rangle. \end{aligned} \tag{4.23}$$

and from (4.22), (4.23) we obtain that for every $0 < t \leq 1$,

$$t\left\langle p, b(x) - \frac{1}{2}\sum_j D\sigma_j(x)\sigma_j(x)\right\rangle = O(t^{3/2}).$$

Dividing by t and taking the limit completes the proof of necessary conditions.

Sufficiency. Fix an $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ a.s. and consider the strong solution $X(t)$ to (2.1).

Set $\psi(t) = \mathbb{E}d_K^2(X(t))$. To prove that $X(t) \in K$ almost surely, we have to show that $\psi(t) = 0$. Since $\psi(0) = 0$, by Proposition 2.7 it is enough to prove that for some $L > 0$ and all $t > 0$ such that $\psi(t) > 0$ we have

$$d\psi(t) \leq L\psi(t).$$

Fix $t > 0$ such that $\psi(t) > 0$, and also an element $\widetilde{X}(t)$ in the class of functions equivalent to $X(t) \in L^2(\Omega; H)$. Set $\varphi_\omega(h) = d_K^2(\widetilde{X}_\omega(h))$.

By the measurable selection theorem [7, p.317] there exists an $\mathcal{F}_t$-measurable map $\omega \mapsto \zeta_\omega \in K$ such that

$$\varphi_\omega(t) = |\widetilde{X}_\omega(t) - \zeta_\omega|^2.$$

In particular, $\widetilde{X}_\omega(t) - \zeta_\omega$ is a proximal normal to K at ζ_ω.

Fix $h \in (0,1)$, and also an element $\widetilde{W}(t)$ in the class of functions equivalent to $W(t) \in L^2(\Omega; H_1)$, and an element $\widetilde{W}(t+h)$ in the class of functions equivalent to $W(t+h) \in L^2(\Omega; H_1)$. For all $\omega \in \Omega$ set

$$u_\omega(h) := \begin{cases} 0 & \text{if } \widetilde{W}_\omega(t+h) = \widetilde{W}_\omega(t) \\ \dfrac{\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)}{|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|} & \text{otherwise.} \end{cases} \tag{4.24}$$

For every $\omega \in \Omega$ consider the solution $z_\omega(\cdot)$ to the deterministic system

$$z'(s) = \sigma(z(s))u_\omega(h), \quad z_\omega(0) = \zeta_\omega.$$

By (4.1) and Proposition 2.5, $z_\omega(s) \in K$ for all $s \geq 0$. On the other hand, by (4.7), for all $\omega \in \Omega$

$$z_\omega(|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|) = \zeta_\omega + \sigma(\zeta_\omega)(\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)) +$$

$$+\frac{1}{2}\left(\sum_j \left\langle \nabla\sigma_{ij}(\zeta_\omega), \sigma(\zeta_\omega)(\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t))\right\rangle (\widetilde{W}^j_\omega(t+h) - \widetilde{W}^j_\omega(t))\right) + \Phi_\omega(t),$$

where for some $M > 0$, $|\Phi_\omega(t)| \leq M|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|^3$. Set

$$y_\omega(h) = z_\omega(|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|).$$

By the same arguments as in the proof of necessary conditions, replacing $\widetilde{X}(t)$ by ζ, we show that $y(h)$ is $\mathcal{F}_{t+h}-$measurable and $y(h) \in L^2(\Omega; H)$. Observe next that

$$d_K^2(X(t+h)) \leq |X(t+h) - y(h)|^2. \tag{4.25}$$

Since

$$X(t+h) = X(t) + \int_t^{t+h} b(X(s))ds + \int_t^{t+h} \sigma(X(s))dW(s) \text{ a.s.},$$

for a constant $\alpha > 1$ independent from t

$$|X(t+h) - y(h)|^2 \le \varphi^2(t) + \alpha h^2 \|b\|_\infty^2 + \alpha \Big| \int_t^{t+h} (\sigma(X(s)) - \sigma(\zeta)) dW(s)\Big|^2 +$$

$$+\alpha \left\| \sum_j \left\langle \nabla\sigma_{ij}(\zeta), \sigma(\zeta)(\widetilde{W}(t+h) - \widetilde{W}(t)) \right\rangle (\widetilde{W}^j(t+h) - \widetilde{W}^j(t)) \right\|^2 + \alpha\Phi(t)^2 +$$

$$+2\left\langle X(t) - \zeta, \int_t^{t+h} b(X(s))ds + \int_t^{t+h} (\sigma(X(s)) - \sigma(\zeta))dW(s) - \Phi(t) \right\rangle -$$

$$-2\left\langle X(t) - \zeta, \frac{1}{2}\left(\sum_j \left\langle \nabla\sigma_{ij}(\zeta), \sigma(\zeta)(\widetilde{W}(t+h) - \widetilde{W}(t)) \right\rangle (\widetilde{W}^j(t+h) - \widetilde{W}^j(t)) \right) \right\rangle$$

$$= J_1(h) + J_2(h) + J_3(h) + J_4(h) + J_5(h) + 2J_6(h) - 2J_7(h).$$

Then $\mathbb{E}(J_2(h)) = O(h^2)$. In the same way as we have shown the estimate of $\mathbb{E}(I_4(t))$ we prove that

$$\mathbb{E}\Big| \int_t^{t+h} (\sigma(X(s)) - \sigma(X(t)))dW(s)\Big|^2 = O(h^2).$$

Thus for a constant $c_6 := 2\alpha$

$$\begin{aligned} \mathbb{E}(J_3(h)) &\le c_6 \mathbb{E}\Big| \int_t^{t+h} (\sigma(X(t)) - \sigma(\zeta))dW(s)\Big|^2 + O(h^2) \\ &= c_6 h \mathbb{E}|\sigma(X(t)) - \sigma(\zeta)|^2 + O(h^2). \end{aligned}$$

Since σ is Lipschitz continuous we deduce that for a constant $c_7 > 0$ independent from t

$$\mathbb{E}(J_3(h)) \le c_7 h \psi(t) + O(h^2). \tag{4.26}$$

Similarly to Proposition 2.1

$$\mathbb{E}|W(t+h) - W(t)|^4 = O(h^2),\ \mathbb{E}|W(t+h) - W(t)|^6 = O(h^3) \tag{4.27}$$

and therefore $\mathbb{E}(J_4(h)) + \mathbb{E}(J_5(h)) = O(h^2)$ for $0 < h \le 1$. By (2.2) and the Hölder inequality, for some $M_0 > 0$ and all $s \in [t, t+1]$,

$$\mathbb{E}|X(s) - X(t)| \le \sqrt{M_0}\sqrt{s-t}.$$

This and the Lipschitz continuity of b imply that for a positive constant c_8 independent from t

$$\mathbb{E}\left\langle X(t)-\zeta, \int_t^{t+h} b(X(s))ds \right\rangle \leq \mathbb{E}\left\langle X(t)-\zeta, hb(\zeta)\right\rangle +$$

$$+\mathbb{E}\left(|X(t)-\zeta| \int_t^{t+h} (|b(X(s))-b(X(t))| + |b(X(t))-b(\zeta)|)ds \right) \leq$$

$$\leq h\mathbb{E}\left\langle X(t)-\zeta, b(\zeta)\right\rangle + hc_8\psi(t) + O(h^{3/2}).$$

Furthermore, $\mathbb{E}\left\langle X(t)-\zeta, \int_t^{t+h}(\sigma(X(s))-\sigma(\zeta))dW(s)\right\rangle = 0$ and, by (4.27) and the Hölder inequality for a constant $c_9 > 0$

$$\mathbb{E}|\left\langle X(t)-\zeta, \Phi(t)\right\rangle| \leq c_9\sqrt{\mathbb{E}|W(t+h)-W(t)|^6} = O(h^{3/2}).$$

Consequently,

$$\mathbb{E}(J_6(h)) \leq h\mathbb{E}\left\langle X(t)-\zeta, b(\zeta)\right\rangle + hc_8\psi(t) + O(h^{3/2}).$$

Finally

$$2\mathbb{E}(J_7(h)) = \mathbb{E}\left\langle X(t)-\zeta, \left(\sum_j \left\langle \nabla\sigma_{ij}(\zeta), \sigma(\zeta)(W(t+h)-W(t))\right\rangle (W^j(t+h)-W^j(t))\right)\right\rangle$$

$$= \mathbb{E}\left(\sum_i (X^i(t)-\zeta^i)\sum_{j,k,r}\frac{\partial\sigma_{ij}}{\partial x_k}(\zeta)\sigma_{kr}(\zeta)(W^r(t+h)-W^r(t))(W^j(t+h)-W^j(t))\right)$$

$$= h\mathbb{E}\left(\sum_i (X^i(t)-\zeta^i)\sum_{j,k}\frac{\partial\sigma_{ij}}{\partial x_k}(\zeta)\sigma_{kj}(\zeta)\right) = h\mathbb{E}\left\langle X(t)-\zeta, \sum_j D\sigma_j(\zeta)\sigma_j(\zeta)\right\rangle.$$

The above inequalities and (4.25) imply that for $L := 2(c_7+c_8)$ and for all $0 < h \leq 1$,

$$\psi(t+h) \leq \psi(t) + Lh\psi(t) + 2h\mathbb{E}\left\langle X(t)-\zeta, b(\zeta) - \frac{1}{2}\sum_j D\sigma_j(\zeta)\sigma_j(\zeta)\right\rangle + O(h^{3/2}).$$

Thus, by (4.1), $\psi(t+h) \leq \psi(t) + Lh\psi(t) + O(h^{3/2})$ and therefore $d\psi(t) \leq L\psi(t)$. By Proposition 2.7, $\psi \equiv 0$ implying that $d_K(X(t)) = 0$ almost surely. □

5 Invariance of Stochastic Control Systems

Let U be a complete separable metric space and $b : H\times U \to H$, and $\sigma : H\times U \to L(H_1, H)$ be bounded continuous mappings. Assume that there exists a constant $C > 0$ such that

$$\forall\, x,\ y \in H,\ \ \forall\, v \in U,\ \ |b(x,v)-b(y,v)| + \|\sigma(x,v)-\sigma(y,v)\| \leq C|x-y|. \tag{5.1}$$

Denote by $\mathcal{A}$ the set of all $L^1(\Omega, U)-$valued mappings $v(\cdot)$ defined on $\mathbb{R}_+$ which are progressively measurable with respect to the family $\mathcal{F}_t$, i.e. for every $t \geq 0$, $v(t) \in U$ a.s. and the mapping $[0,t] \times \Omega \ni (s,\omega) \mapsto v_\omega(s)$ is $\mathcal{B}_1 \times \mathcal{F}_t-$measurable. Elements of $\mathcal{A}$ are called admissible controls.

We associate to the above data the stochastic control system

$$dX = b(X, v(t))dt + \sigma(X, v(t))dW(t), \quad v(\cdot) \in \mathcal{A}. \tag{5.2}$$

Let $X_0 \in L^2(\Omega)$ be an $\mathcal{F}_0$-random variable, $v(\cdot) \in \mathcal{A}$ and consider the differential stochastic equation

$$\begin{cases} dX = b(X, v(t))dt + \sigma(X, v(t))dW(t), \\ \\ X(0) = X_0. \end{cases} \tag{5.3}$$

Under the above assumptions (5.3) has a unique solution $X(\cdot)$, i.e. for all $t \geq 0$,

$$X(t) = X_0 + \int_0^t b(X(s), v(s))ds + \int_0^t \sigma(X(s), v(s))dW(s) \ \text{ a.s..}$$

(see [27, Chapters 1, 2]).

Definition 5.1 *A set $K \subset H$ is called invariant under the control system (5.2) if for every $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ almost surely and every admissible control $v(\cdot) \in \mathcal{A}$, the solution X to (5.3) satisfies for all $t \geq 0$, $X(t) \in K$ almost surely.*

Theorem 5.2 *Assume that K is closed, b, σ are bounded and continuous, that there exists a constant $C > 0$ such that (5.1) holds true and for all $v \in U$, $\sigma'(\cdot, v)$ is $C-$Lipschitz. Then K is invariant under (5.2) if and only if for every $x \in \partial K$ and for all proximal normal p to K at x we have*

$$\left\langle p, \ b(x,v) - \frac{1}{2}\sum_{j=1}^m D_x\sigma_j(x,v)\sigma_j(x,v) \right\rangle \leq 0, \quad \sigma(x,v)^*p = 0, \quad \forall\, v \in U, \tag{5.4}$$

where $\sigma_j(x,v)$ denotes the column j of the matrix $\sigma(x,v)$ and $D_x\sigma_j(x,v)$ the jacobian of $\sigma_j(\cdot, v)$ at x.

Remark 5.3 Exactly as in Remark 4.2 and Corollary 4.3 in the above theorem proximal normals may be replaced by the elements of normal cone $N_K(x)$ or by those of Clarke's normal cone.

Corollary 5.4 *If all the assumptions of Theorem 5.2 hold true, then K is invariant under (5.2) if and only if K is invariant under the deterministic control system with two controls*

$$\begin{cases} x' = b(x, v(t)) - \frac{1}{2}\sum_{j=1}^m D\sigma_j(x, v(t))\sigma_j(x, v(t)) + \sigma(x, v(t))u(t), \\ \\ u \in L^1_{loc}(\mathbb{R}_+, H_1), \quad v : \mathbb{R}_+ \to U \ \text{ is measurable} \end{cases} \tag{5.5}$$

or, equivalently, if and only if

$$b(x,v) - \frac{1}{2}\sum_{j=1}^m D\sigma_j(x,v)\sigma_j(x,v) \in T_K(x), \ \mathrm{Im}(\sigma(x,v)) \subset T_K(x), \quad \forall\, v \in U, \ \forall\, x \in K.$$

Proof — By Remark 5.3, (5.4) holds true, if and only if it holds true for all $p \in N_K(x)$. By the separation theorem, relations (5.4) may be equivalently written as

$$\forall\, x \in K,\ \forall\, v\ \in U,\ b(x,v) - \frac{1}{2}\sum_{j=1}^{m} D\sigma_j(x,v)\sigma_j(x,v) \in \overline{\mathrm{co}}\, T_K(x)\ \&\ \mathrm{Im}(\sigma(x,v)) \subset \overline{\mathrm{co}}\, T_K(x).$$

By [7, Theorem 4.1.10] and continuity of b, σ, in the above $\overline{\mathrm{co}}\, T_K(x)$ may be replaced by $T_K(x)$. Hence, by [1], (5.4) is a necessary and sufficient condition for the invariance of K under the deterministic control system (5.5). $\square$

Theorem 5.2 and Lemma 2.3 imply the following result.

Corollary 5.5 *Under the assumptions of Theorem 5.2 suppose that for every $x \in \partial K$, $p \in N_K(x)$ and $v \in U$, $DN_K(x,p)(\sigma_j(x,v)) \neq \emptyset$ for all $j = 1, ..., m$. Then the set K is invariant under the system (5.2) if and only if for every $x \in \partial K$ and $p \in N_K(x)$*

$$\langle p, b(x,v)\rangle + \frac{1}{2}\sum_{j=1}^{m} \mathrm{Curv}_K(x,p)(\sigma_j(x,v),\sigma_j(x,v)) \le 0,\ \sigma(x,v)^* p = 0,\ \forall\, v \in U. \tag{5.6}$$

Furthermore, for all $v \in U$ and $\mu_j \in DN_K(x,p)(\sigma_j(x,v))$

$$\mathrm{Curv}_K(x,p)(\sigma_j(x,v),\sigma_j(x,v)) = \langle \mu_j, \sigma_j(x,v)\rangle = -\langle p, D\sigma_j(x,v)\sigma_j(x,v)\rangle\,.$$

Proof of Theorem 5.2 — If the set K is invariant under the system (5.2), then for every $v_0 \in U$ the mapping $v \equiv v_0$ belongs to $\mathcal{A}$. Thus, for every $\mathcal{F}_0$-measurable random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ a.s., the solution X to (5.3) satisfies $X(t) \in K$ a.s. This and Theorem 4.1 imply that for every $x \in \partial K$ and for all proximal normal p to K at x relations (5.4) hold true.

Assume next that (5.4) holds true for every $x \in \partial K$ and any proximal normal p to K at x. To prove the invariance we proceed in several steps. We first show that (5.4) implies the invariance for time independent controls, then for piecewise constant (with respect to the time) controls and, finally, in the general case.

Case 1 of constant controls. The proof is essentially the same as the one of the invariance of differential stochastic equations, but the setting is slightly different because of the presence of the control v. Let $t_0 \ge 0$. Consider an $\mathcal{F}_{t_0}$-measurable $v : \Omega \ni \omega \mapsto U$, an $\mathcal{F}_{t_0}$-random variable $X_0 \in K$ a.s. and the strong solution $X(t)$ to

$$dX = b(X,v)dt + \sigma(X,v)dW(t),\quad X(t_0) = X_0,\quad t \ge t_0.$$

Define $\psi(t) = \mathbb{E} d_K^2(X(t))$. Then $\psi(t_0) = 0$. As in the proof of sufficient conditions for stochastic differential equations, we check that for some $L > 0$ and all $t > t_0$ with $\psi(t) > 0$, we have $d\psi(t) \le L\psi(t)$. Fix such $t > t_0$ and an element $\widetilde{X}(t)$ in the class of functions equivalent to $X(t) \in L^2(\Omega; H)$. Set $\varphi_\omega(t) = d_K^2(\widetilde{X}_\omega(t))$.

Consider an $\mathcal{F}_t$-measurable map $\omega \mapsto \zeta_\omega \in K$ satisfying $\varphi_\omega(t) = |\widetilde{X}_\omega(t) - \zeta_\omega|^2$.

Fix $h \in (0,1)$, and also an element $\widetilde{W}(t)$ in the class of functions equivalent to $W(t) \in L^2(\Omega; H_1)$, an element $\widetilde{W}(t+h)$ in the class of functions equivalent to $W(t+h) \in L^2(\Omega; H_1)$.

Define $u_\omega(h)$ as in (4.24). For every $\omega \in \Omega$, consider the solution $z_\omega(\cdot)$ to the deterministic system

$$z'(s) = \sigma(z(s), v_\omega)u_\omega(h), \quad z_\omega(0) = \zeta_\omega.$$

By (5.4) and Proposition 2.5, $z_\omega(s) \in K$ for all $s \geq 0$.

On the other hand, by (4.7), for all $\omega \in \Omega$

$$z_\omega(|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|) = \zeta_\omega + \sigma(\zeta_\omega, v_\omega)(\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)) +$$

$$+\frac{1}{2}\left(\sum_j \left\langle \nabla_x \sigma_{ij}(\zeta_\omega, v_\omega), \sigma(\zeta_\omega, v_\omega)(\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t))\right\rangle (\widetilde{W}^j_\omega(t+h) - \widetilde{W}^j_\omega(t))\right) + \Phi_\omega(t),$$

where for some $M > 0$, $|\Phi_\omega(t)| \leq M|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|^3$. Set

$$y_\omega(h) = z_\omega(|\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|).$$

We claim that $y(h)$ is $\mathcal{F}_{t+h}$-measurable and $y(h) \in L^2(\Omega, H)$. Indeed let $(x_0, u_0, v_0) \in H \times H_1 \times U$ and consider the solution $z(\cdot; x_0, u_0, v_0)$ to $z' = \sigma(z, v_0)u_0$ satisfying $z(0) = x_0$. Define the closed set

$$\Pi := \{(s, x_0, u_0, v_0, z(s; x_0, u_0, v_0)) \mid s \geq 0,\ x_0 \in H,\ u_0 \in H_1,\ v_0 \in U\}$$

and the set $\Gamma(h) := \{(\omega, |\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|, \zeta_\omega, u_\omega, v_\omega, r) \mid \omega \in \Omega,\ r \in H\} \in \mathcal{F}_{t+h} \times \mathcal{B}_1 \times \mathcal{B}_H \times \mathcal{B}_{H_1} \times \mathcal{B}_U \times \mathcal{B}_H$, where $\mathcal{B}_U$ denotes the $\sigma-$ algebra of Borel subsets of U. Then $\Gamma(h) \cap (\Omega \times \Pi) \in \mathcal{F}_{t+h} \times \mathcal{B}_1 \times \mathcal{B}_H \times \mathcal{B}_{H_1} \times \mathcal{B}_U \times \mathcal{B}_H$. Since

$$\Gamma(h) \cap (\Omega \times \Pi) = \{(\omega, |\widetilde{W}_\omega(t+h) - \widetilde{W}_\omega(t)|, \zeta_\omega, u_\omega, v_\omega, y_\omega(h)) \mid \omega \in \Omega\},$$

by the projection theorem (see for instance [11]), $\omega \mapsto y_\omega(h)$ is $\mathcal{F}_{t+h}$-measurable. Since σ is bounded, we deduce that $y(h) \in L^2(\Omega, H)$. Notice that $\widetilde{X}_\omega(t) - \zeta_\omega$ is a proximal normal to K at ζ_ω and (4.25) holds true. By exactly the same arguments as those used in the proof of sufficiency of Theorem 4.1, we check that for some $L > 0$ independent from t and all $h \in [0,1]$, $\psi(t+h) \leq \psi(t) + Lh\psi(t) + O(h^{3/2})$ and therefore $d\psi(t) \leq L\psi(t)$. By Proposition 2.7, $\psi \equiv 0$ implying that $d_K(X(t)) = 0$ almost surely for $t \geq t_0$.

Case 2 of piecewise constant controls. Let $v \in \mathcal{A}$ be such that for some $0 = s_0 < s_1 < ... < s_k < ...$ and for all $k \geq 0$, v is time independent on the time interval $[s_k, s_{k+1})$. Fix an $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ a.s. and consider the solution $X(s)$ to

$$dX = b(X, v)dt + \sigma(X, v)dW(t), \quad X(0) = X_0, \quad t \in [s_0, s_1).$$

Then, by Case 1, for all $0 \leq s < s_1$, $X(s) \in K$ a.s. Since b, σ are bounded, X can be extended by continuity to s_1 and $X(s_1) \in K$ a.s. Assume, that we already proved that for some $k \geq 1$, $X(s) \in K$ a.s. for all $s \leq s_k$. Set $X_0 := X(s_k)$ and consider the solution $X(t)$ to

$$dX = b(X, v)dt + \sigma(X, v)dW(t), \quad X(s_k) = X_0, \quad t \in [s_k, s_{k+1}).$$

By Case 1, for all $s \in [s_k, s_{k+1})$, $X(s) \in K$ a.s. and again we extend X by continuity to s_{k+1}. This and the induction argument yield that $X(s) \in K$ a.s. for all $s \geq 0$.

The general case. Consider $v \in \mathcal{A}$, $\mathcal{F}_0$-random variable $X_0 \in L^2(\Omega)$ such that $X_0 \in K$ a.s. and the solution $X(t)$ to (5.3) (see for instance [27, p. 42]).

Fix $t > 0$. Then for some $M_0 > 0$, the inequality (2.2) holds true with $t_0 = t$. We have to show that $X(t) \in K$ a.s. For this end let us fix $0 < \varepsilon < 1$ and define the mapping

$$\mathbb{R}_+ \ni s \mapsto g(s) := b(X(s), v(s)) \in L^2(\Omega, H).$$

Then for all $\omega \in \Omega$, the mapping $s \mapsto \int_0^s g_\omega(\tau)d\tau \in H$ is absolutely continuous on bounded intervals. Define the mapping $f : [0,t] \times (\mathbb{R}_+ \backslash \{0\}) \to L^2(\Omega, H)$ by

$$f(s,h) = \frac{1}{h}\int_s^{s+h} g(\tau)d\tau.$$

Then, by the absolute continuity and boundedness of b, for all $\omega \in \Omega$,

$$\lim_{h\to 0+} \int_0^t |f_\omega(s,h) - g_\omega(s)|ds = 0.$$

Hence, by the dominated convergence theorem,

$$\lim_{h\to 0+} \mathbb{E}\int_0^t |f(s,h) - g(s)|ds = 0.$$

Next, applying the Fubini theorem, we obtain that

$$\lim_{h\to 0+} \int_0^t \mathbb{E}|f(s,h) - g(s)|ds = 0. \tag{5.7}$$

Let $h_i \to 0+$ be such for all $i \geq 1$, $h_i \leq \varepsilon^2$.

Claim. We claim that for all i large enough, there exist $\delta_i \to 0+$ and $0 = \tau_0^i \leq s_1^i < \tau_1^i \leq s_2^i ... \leq s_{m_i}^i < \tau_{m_i}^i \leq t$ such that for all $1 \leq j \leq m_i$, $\tau_j^i = s_j^i + h_i$ and

$$0 \leq t - m_i h_i \leq \delta_i + h_i, \quad \mathbb{E}|\int_{s_j^i}^{s_j^i+h_i} g(\tau)d\tau - h_i g(s_j^i)| \leq \varepsilon h_i. \tag{5.8}$$

Indeed, define the measurable sets $A_i := \{s \in [0,t] \mid \mathbb{E}|f(s,h_i) - g(s)| > \varepsilon\}$. By (5.7) the Lebesgue measures $\mu(A_i)$ converge to zero when $i \to \infty$. Set $\delta_i = \mu(A_i) + 1/i$. Consider open in $[0,t]$ sets $\mathcal{O}_i$ such that $A_i \subset \mathcal{O}_i$ and $\mu(\mathcal{O}_i) \leq \delta_i$. Then the sets $C_i := [0,t] \backslash \mathcal{O}_i$ are closed subsets of $[0,t]$. Fix $i \geq 1$ such that there exists $s \in C_i$ with $s + h_i \leq t$.

Set $s_1^i := \min\{s \mid s \in C_i\}$ and $\tau_1^i := s_1^i + h_i$. Inductively, assume that we already constructed for some $k \geq 1$, the numbers $0 = \tau_0^i \leq s_1^i < \tau_1^i \leq ... < \tau_k^i \leq t$ such that for all $1 \leq j \leq k$, $\tau_j^i - s_j^i = h_i$, $s_j^i \in C_i$ and $[\tau_{j-1}^i, s_j^i) \subset \mathcal{O}_i$.

If $C_i \cap [\tau_k^i, t] = \emptyset$, then put $m_i = k$. If $C_i \cap [\tau_k^i, t] \neq \emptyset$, then define

$$s_{k+1}^i := \min\{s \mid s \in C_i \cap [\tau_k^i, t]\}.$$

Clearly $[\tau_k^i, s_{k+1}^i) \subset \mathcal{O}_i$. If $s_{k+1}^i + h_i \le t$, then define $\tau_{k+1}^i := s_{k+1}^i + h_i$, otherwise put $m_i = k$.

Since the interval $[0, t]$ is finite and $h_i > 0$ is fixed, this construction ends in a finite number of steps. In this way we defined also m_i. On the other hand, for all $0 \le j \le m_i - 1$, $[\tau_j^i, s_{j+1}^i) \subset \mathcal{O}_i$. Furthermore, by our construction, either $[\tau_{m_i}^i, t] \subset \mathcal{O}_i$, or there exists $s_{m_i+1}^i$ such that $[\tau_{m_i}^i, s_{m_i+1}^i) \subset \mathcal{O}_i$ and $t - s_{m_i+1}^i < h_i$. Since

$$\sum_{j=1}^{m_i} (\tau_j^i - s_j^i) + \sum_{j=0}^{m_i-1} (s_{j+1}^i - \tau_j^i) + (t - \tau_{m_i}) = t,$$

we have

$$t - m_i h_i = \sum_{j=0}^{m_i-1} (s_{j+1}^i - \tau_j^i) + (t - \tau_{m_i}) \le \delta_i + h_i$$

and our claim is proved.

Define piecewise constant controls

$$u_i(s) := \begin{cases} v(s_j^i) & \text{if for some } 1 \le j \le m_i - 1, \;\; s \in [s_j^i, \tau_j^i), \\ v(\tau_j^i) & \text{if for some } 0 \le j \le m_i - 1, \;\; s \in [\tau_j^i, s_{j+1}^i), \\ v(\tau_{m_i}^i) & \text{if } s \ge \tau_{m_i}^i, \end{cases}$$

and piecewise constant functions

$$X_i(s) := \begin{cases} X(s_j^i) & \text{if for some } 1 \le j \le m_i - 1, \;\; s \in [s_j^i, \tau_j^i), \\ X(\tau_j^i) & \text{if for some } 0 \le j \le m_i - 1, \;\; s \in [\tau_j^i, s_{j+1}^i) \\ X(\tau_{m_i}^i) & \text{if } s \ge \tau_{m_i}^i. \end{cases}$$

Then, be the very definition of the Itô integral, for all $s \in [0, t]$,

$$\lim_{i \to \infty} \mathbb{E}| \int_0^s (\sigma(X(\rho), u(\rho)) - \sigma(X_i(\rho), u_i(\rho))) dW(\rho)|^2 = 0. \tag{5.9}$$

Consider solutions Y_i to

$$dY = b(Y, u_i)dt + \sigma(Y, u_i)dW(t), \;\; X(0) = X_0.$$

Then by the Case 2, $Y_i(s) \in K$ a.s. for all $s \ge 0$. Set $\psi_i^\varepsilon(s) := \mathbb{E}|X(s) - Y_i(s)|^2$ and notice that

$$\mathbb{E} d_K^2(X(t)) \le \psi_i^\varepsilon(t), \tag{5.10}$$

and that for some $\alpha > 1$ and all $s \in [0,t]$,

$$\begin{aligned}
\frac{1}{\alpha}\psi_i^\varepsilon(s) \ &\leq \int_0^s \mathbb{E}|b(X(\rho),u_i(\rho)) - b(Y_i(\rho),u_i(\rho))|^2 d\rho + \\
&+\mathbb{E}|\int_0^s (\sigma(X(\rho),u_i(\rho)) - \sigma(Y_i(\rho),u_i(\rho)))dW(\rho)|^2 + \\
&+\mathbb{E}|\int_0^s (\sigma(X(\rho),u_i(\rho)) - \sigma(X_i(\rho),u_i(\rho)))dW(\rho)|^2 + \\
&+\mathbb{E}|\int_0^s (g(\rho) - b(X(\rho),u_i(\rho))d\rho|^2 + \\
&+\mathbb{E}|\int_0^s (\sigma(X(\rho),u(\rho)) - \sigma(X_i(\rho),u_i(\rho)))dW(\rho)|^2 = \\
&= I_1^i(s) + I_2^i(s) + I_3^i(s) + I_4^i(s) + I_5^i(s).
\end{aligned} \tag{5.11}$$

Since b and σ are $C-$Lipschitz in the first variable,

$$I_1^i(s) + I_2^i(s) \leq 2C\int_0^s \mathbb{E}|X(\rho) - Y_i(\rho)|^2 d\rho = 2C\int_0^s \psi_i^\varepsilon(\rho)d\rho. \tag{5.12}$$

By the Lipschitz continuity of σ with respect to x, using that $\mathbb{E}|X(\rho)|^2 \leq Mt$ for all $\rho \in [0,t]$, for a constant $C_1 > 0$ independent from $s \in [0,t]$

$$\begin{aligned}
&I_3^i(s) \leq C\int_0^s \mathbb{E}|X(\rho) - X_i(\rho)|^2 d\rho \leq C_1\sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}(\rho - s_j^i)d\rho + 2MCt(\delta_i + h_i) \leq \\
&\leq C_1\sum_{j=1}^{m_i}(\tau_j^i - s_j^i)^2 + 2MCt(\delta_i + h_i) \leq C_1\varepsilon^2 t + 2MCt(\delta_i + \varepsilon^2).
\end{aligned} \tag{5.13}$$

On the other hand

$$\begin{aligned}
&\left|\int_0^s (g(\rho) - b(X(\rho),u_i(\rho)))d\rho\right| \leq \sum_{j=1}^{m_i}|\int_{s_j^i}^{\tau_j^i}(g(\rho) - b(X(s_j^i),v(s_j^i)))d\rho| + \\
&+2\|b\|_\infty(\varepsilon^2 + \delta_i) + \sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}|b(X(\rho),u_i(\rho)) - b(X(s_j^i),v(s_j^i))|d\rho.
\end{aligned} \tag{5.14}$$

By (5.8)

$$\mathbb{E}\left(\sum_{j=1}^{m_i}|\int_{s_j^i}^{\tau_j^i}(g(\rho) - b(X(s_j^i),v(s_j^i)))d\rho|\right) \leq \varepsilon t,$$

which implies that

$$\mathbb{E}\left(\sum_{j=1}^{m_i}|\int_{s_j^i}^{\tau_j^i}(g(\rho)-b(X(s_j^i),v(s_j^i)))d\rho|\right)^2 \le 2\varepsilon t^2\|b\|_\infty. \tag{5.15}$$

Furthermore, by the Lipschitz continuity of b with respect to x and by (2.2), for a constant $c_1>0$ independent from i, ε

$$\begin{aligned}
&\mathbb{E}\left(\sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}|b(X(\rho),u_i(\rho))-b(X(s_j^i),v(s_j^i))|d\rho\right)^2 \le \\
&\le 2t\|b\|_\infty\sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}\mathbb{E}|b(X(\rho),u_i(\rho))-b(X(s_j^i),v(s_j^i))|d\rho \le \\
&\le 2Ct\|b\|_\infty\sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}\mathbb{E}|X(\rho)-X(s_j^i)|d\rho \le c_1\sum_{j=1}^{m_i}\int_{s_j^i}^{\tau_j^i}\sqrt{\rho-s_j^i}\;d\rho \le \\
&\le c_1\sum_{j=1}^{m_i}(\tau_j^i-s_j^i)^{3/2}\le c_1\varepsilon t.
\end{aligned} \tag{5.16}$$

From (5.14) - (5.16) we deduce that for constant $c_2>0$ independent from i, ε and s

$$I_4^i(s)\le c_2(\varepsilon+\delta_i).$$

This and (5.11) - (5.13) imply for a constant $c_3>0$ independent from i, ε and for all $s\in[0,t]$

$$\psi_i^\varepsilon(s)\le c_3\textstyle\int_0^s\psi_i^\varepsilon(\rho)d\rho+c_3(\varepsilon+\delta_i)+I_5^i(s).$$

Then it follows from the Gronwall inequality that for a constant $c_4>0$ independent from ε and for all i,

$$\psi_i^\varepsilon(t)\le c_4(\varepsilon+\delta_i+I_5^i(t))+c_4\textstyle\int_0^t I_5^i(s)ds. \tag{5.17}$$

From (5.9) we know that for every $s\in[0,t]$, $\lim_{i\to\infty}I_5^i(s)=0$. On the other hand,

$$I_5^i(s)=\int_0^s\mathbb{E}\|\sigma(X(\rho),u(\rho))-\sigma(X_i(\rho),u_i(\rho))\|^2d\rho\le 2s\|\sigma\|_\infty^2.$$

From the Lebesgue dominated convergence theorem we deduce that

$$\lim_{i\to\infty}\int_0^t I_5^i(s)ds=0.$$

This and (5.10), (5.17) imply that $\mathbb{E}d_K^2(X(t)\le\limsup_{i\to\infty}\psi_i^\varepsilon(t)\le c_4\varepsilon$. Since $\varepsilon>0$ is arbitrary and c_4 does not depend on ε, we get $X(t)\in K$ a.s. $\square$

Remark 5.6 Recall that the second order normal cone $N_K^2(x)$ to K at x is the set of all $(p,Q) \in H \times L(H,H)$ satisfying

$$\forall\, y \in K, \quad \langle p, y-x\rangle + \frac{1}{2}Q(y-x,y-x) \le o(|y-x|^2).$$

It is not difficult to realize that if $(p,Q) \in N_K^2(x)$, then $p \in N_K(x)$.

If K is invariant under the stochastic control system (5.2), then for every $x \in \partial K$, and for all $p \in N_K(x)$ relations (5.4) hold true. Fix $x \in K,\ v \in U$. Applying Proposition 2.5 with $u(t) \equiv e_j$ and σ replaced by $\sigma(\cdot,v)$we deduce that for all $h > 0$,

$$y_h := x + \sqrt{h}\sigma_j(x,v) + \frac{h}{2}D\sigma_j(x,v)\sigma_j(x,v) + o(h) \in K.$$

Using that $\langle p, \sigma_j(x,v)\rangle = 0$, from the definition of second order normals it follows that for all $(p,Q) \in N_K^2(x)$,

$$\frac{h}{2}\langle p, D\sigma_j(x,v)\sigma_j(x,v)\rangle + \frac{h}{2}Q(\sigma_j(x,v),\sigma_j(x,v)) \le o(h).$$

Dividing by h and taking the limit yields $Q(\sigma_j(x,v),\sigma_j(x,v)) \le -\langle p, D\sigma_j(x,v)\sigma_j(x,v)\rangle$. This and (5.4) imply that

$$\forall\, (p,Q) \in N_K^2(x), \quad \langle p, b(x,v)\rangle + \frac{1}{2}\mathrm{Tr}[Q\sigma(x,v)\sigma^*(x,v)] \le 0, \tag{5.18}$$

i.e. (5.4) yields a necessary condition for the invariance of stochastic control systems proposed in [8]. We also observe that (5.18) is a simple consequence of the Itô formulae and the definition of $N_K^2(x)$. So (5.4) is not really needed to prove that (5.18) is a necessary condition for the invariance.

The difference in the presentation of sufficient conditions seems to be however important. Namely in [8] it is also proved, using the viscosity solutions approach, that the second order condition (5.18) is sufficient for the invariance when the initial conditions are deterministic (i.e. are elements of H). Since there is no calculus available for the second order normal cones, it is not clear how to deduce directly from the second order condition (5.18) our first order conditions (5.4), except in the case of smooth boundaries ∂K.

References

[1] AUBIN J.-P. (1991) VIABILITY THEORY, Birkhäuser, Boston, Basel, Berlin.

[2] AUBIN J.-P. & DA PRATO G. (1990) *Stochastic viability and invariance,* Annali Scuola Normale di Pisa, 27, 595-694.

[3] AUBIN J.-P. & DA PRATO G. (1995) *Stochastic Nagumo's viability theorem*, Stochastic Analysis and Applictions, 13, 1-11.

[4] AUBIN J.-P. & DA PRATO G. (1998) *The viability theorem for stochastic differential inclusions*, Stochastic Analysis and Applications, 16, 1-15.

[5] AUBIN J.-P., DA PRATO G. & FRANKOWSKA H. (2000) *Stochastic invariance for differential inclusions*, Set-Valued Analysis, 8, 181-201.

[6] AUBIN J.-P. & DOSS H. (2003) *Characterization of stochastic viability of any nonsmooth set involving its generalized contingent curvature*, Stochastic Analysis and Applications, 21, 955-981.

[7] AUBIN J.-P. & FRANKOWSKA H. (1990) SET-VALUED ANALYSIS, Birkhäuser, Boston, Basel, Berlin.

[8] BARDI M. & GOATIN P. (1999) *Invariant sets for controlled degenerate diffusions: a viscosity solutions approach,* W. M. M Mc Eneaney, G. G. Yin, Q. Zhang Eds. LNiM 1660, Birkhäuser, 191–208.

[9] BARDI M. & JENSEN R. (2001) *A geometric characterization of viable sets for controlled degenerate diffusions*, preprint.

[10] BUCKDAHN R., PENG S., QUINCAMPOIX M. & RAINER C. (1998) *Existence of stochastic control under state constraints*, Comptes-Rendus de l'Académie des Sciences, 327, 17-22.

[11] CASTAING C. & VALADIER M. (1977) CONVEX ANALYSIS AND MEASURABLE MUTIFUNCTIONS, Springer-Verlag, Berlin.

[12] DA PRATO G. (1998) INTRODUCTION TO DIFFERENTIAL STOCHASTIC EQUATIONS, Appunti dei corsi, Scuola Normale di Pisa.

[13] DA PRATO G. & FRANKOWSKA H. (1994) *A stochastic Filippov theorem*, Stochastic Calculus, 12, 409-426.

[14] DA PRATO G. & FRANKOWSKA H. (2001) *Stochastic viability for compact sets in terms of the distance function,* Dynamic Systems and Applications, 10, 177-184.

[15] DOSS H. (1977) *Liens entre équations différentielles stochastiques et ordinaires*, Ann. Inst. Henri Poincaré, Calcul des Probabilités et Statistique, 23, 99-125.

[16] DOSS H. & LENGLART E. (1977) *Sur le comportement asymptotique des solutions d'équations différentielles stochastiques*, CRAS, 284, 971-974.

[17] FRIEDMAN A. (1976) STOCHASTIC DIFFERENTIAL EQUATIONS AND APPLICATIONS, vol.2, Academic Press, New York.

[18] GAUTIER S. & THIBAULT L. (1993) *Viability for constrained stochastic differential equations*, Differential Integral Equations, 6, 1395-1414.

[19] IKEDA N. & WATANABE S. (1981) STOCHASTIC DIFFERENTIAL EQUATIONS AND DIFFUSION PROCESSES, North-Holland.

[20] JAKIMIAK W. (1996) *A note on invariance for semilinear differential equations,* Bull. Pol. Sci., 179–183.

[21] KALLIANPUR G. (1980) STOCHASTIC FILTERING, Springer Verlag, New York

[22] KISIELEWICZ M. (1995) *Viability theorem for stochastic inclusions*, Discussiones Mathematicae, Differential Inclusions, 15, 61-74.

[23] MILIAN A. (1993) *A note on the stochastic invariance for Itô equations, Bull. Polish Acad. Sciences,* 41, 139-150.

[24] MILIAN A. (1995) *Stochastic viability and a comparison theorem*, Colloq. Math., 68, 297-316.

[25] MILIAN A. (1997) *Invariance for stochastic equations with regular coefficients*, Stochastic Anal. Appl., 15, 91-101.

[26] STROOCK D. V. & VARHADAN S. R. S. (1979) MULTIDIMENSIONAL DIFFUSION PROCESSES, Springer–Verlag.

[27] YONG J. & ZHOU X. Y. (1999) STOCHASTIC CONTROLS, Springer–Verlag.